GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 25, 2005, 20:19:56 ; Search time 75 Seconds
(without alignments)
301.479 Million cell updates/sec

Title: US-09-874-140-2
Perfect score: 1206
Sequence: 1 MTQPASSRVVFDPSNPTTFS.....KGQTLSHTGQSGNASRSRRV 235

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1206	100.0	235	1	WZBE23	gene 23 protein -
2	154.5	12.8	119	2	T42568	probable capsid pr
3	153.5	12.7	119	1	WZBEB7	13.6K capsid prote
4	139	11.5	710	1	QQBE22	membrane antigen g
5	131.5	10.9	2232	2	T34434	hypothetical prote
6	128	10.6	310	1	PIHUSD	salivary proline-r
7	128	10.6	1151	2	T18535	high molecular mas
8	127	10.5	721	2	E70766	hypothetical prote
9	127	10.5	780	2	T00366	hypothetical prote
10	126	10.4	775	2	I49759	hepatocyte growth
11	125	10.4	1357	2	T29265	hypothetical prote
12	124.5	10.3	907	1	QQBE21	membrane antigen g
13	124.5	10.3	1032	2	T34433	hypothetical prote
14	123	10.2	392	1	PIHUB6	salivary proline-r
15	122.5	10.2	886	2	S29605	glycoprotein 350/2
16	122.5	10.2	1275	2	T33369	hypothetical prote
17	122	10.1	1096	1	S61917	protein kinase C (
18	122	10.1	1199	2	A40670	nuclear envelope p
19	122	10.1	2441	2	S39161	CREB-binding prote
20	121.5	10.1	3164	1	WMBEH6	UL36 protein - hum
21	121	10.0	251	1	PIHUPF	salivary proline-r
22	121	10.0	678	2	JC4245	transcription fact
23	119.5	9.9	604	2	A39369	homeotic protein B
24	119	9.9	1415	1	EDBEGA	immediate-early pr
25	118.5	9.8	542	2	S64030	probable membrane
26	118	9.8	319	2	T11592	hypothetical prote
27	118	9.8	333	2	S11484	CABP1 protein - sl
28	117.5	9.7	220	2	A36298	proline-rich prote
29	117.5	9.7	1794	2	T38459	hypothetical diver
30	117	9.7	606	2	S13367	Om(1D) protein - f
31	117	9.7	1952	2	T48814	hypothetical prote
32	117	9.7	2440	2	S39162	transcription coac
33	116.5	9.7	677	2	T00369	hypothetical prote
34	115.5	9.6	419	2	T04530	hypothetical prote
35	115.5	9.6	419	2	D85334	hypothetical prote
36	115.5	9.6	838	1	EEWTHW	glutenin, high mol
37	115.5	9.6	1390	2	T31353	polyprotein - Arab
38	115.5	9.6	3942	2	T42730	Bassoon protein -
39	115	9.5	666	2	B70803	hypothetical prote
40	114.5	9.5	212	2	B36298	proline-rich prote
41	114.5	9.5	324	2	S13497	cAMP-binding prote
42	114.5	9.5	510	2	H84824	En/Spm-like transp
43	114.5	9.5	539	2	T28770	hypothetical prote
44	114.5	9.5	587	2	T41653	probable transcrip
45	113.5	9.4	309	2	S10889	proline-rich prote

ALIGNMENTS

RESULT 1
WZBE23
gene 23 protein - human herpesvirus 3
C;Species: human herpesvirus 3, varicella-zoster virus
C;Date: 30-Sep-1988 #sequence_revision 30-Sep-1988 #text_change 09-Jul-2004
C;Accession: E27343
R;Davison, A.J.; Scott, J.E.
J. Gen. Virol. 67, 1759-1816, 1986
A;Title: The complete DNA sequence of varicella-zoster virus.
A;Reference number: A27345; MUID:86306657; PMID:3018124
A;Accession: E27343
A;Molecule type: DNA
A;Residues: 1-235 <DAV>
A;Cross-references: UNIPROT:P09279; EMBL:X04370; NID:g59989; PIDN:CAA27906.1; PID:g60012
C;Genetics:
A;Gene: 23
C;Superfamily: varicella-zoster virus gene 23 protein; human herpesvirus 1 UL35 protein h
F;16-104/Domain: human herpesvirus 1 UL35 protein homology <UL35>

Query Match 100.0%; Score 1206; DB 1; Length 235;
Best Local Similarity 100.0%; Pred. No. 2.7e-69;
Matches 235; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MTQPASSRVVFDPSNPTTFSVEAIAAYTPVALIRLLNASGPLQPGHRVDIADARSIYTVG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MTQPASSRVVFDPSNPTTFSVEAIAAYTPVALIRLLNASGPLQPGHRVDIADARSIYTVG 60

Qy         61 AAASAARARANHNANTIRRTAMFAETDPMTWLRPTVGLKRTFNPRIIRPQPPNPSMSLGI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AAASAARARANHNANTIRRTAMFAETDPMTWLRPTVGLKRTFNPRIIRPQPPNPSMSLGI 120

Qy        121 SGPTILPQKTQSADQSALQQPAALAFSGSSPQHPPPQTTSASVGQQQHVVSGSSGQQPQQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SGPTILPQKTQSADQSALQQPAALAFSGSSPQHPPPQTTSASVGQQQHVVSGSSGQQPQQ 180

Qy        181 GAQSSTVQPTTGSPPAAQGVPQSTPPPTQNTPQGGKGQTLSHTGQSGNASRSRRV 235
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GAQSSTVQPTTGSPPAAQGVPQSTPPPTQNTPQGGKGQTLSHTGQSGNASRSRRV 235
```

RESULT 2
T42568
probable capsid protein 25 - equine herpesvirus 4 (strain NS80567)
C;Species: equine herpesvirus 4
A;Variety: strain NS80567
C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jul-2004
C;Accession: T42568
R;Telford, E.A.; Watson, M.S.; Perry, J.; Cullinane, A.A.; Davison, A.J.
J. Gen. Virol. 79, 1197-1203, 1998
A;Title: The DNA sequence of equine herpesvirus-4.
A;Reference number: Z22173; MUID:98264497; PMID:9603335
A;Accession: T42568

```
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-119 <TEL>
A;Cross-references: UNIPROT:O39267; EMBL:AF030027; NID:g2605950; PIDN:AAC59540.1; PID:g2
A;Experimental source: strain NS80567
C;Genetics:
A;Note: 25
C;Superfamily: human herpesvirus 1 UL35 protein; human herpesvirus 1 UL35 protein homolo

  Query Match             12.8%;  Score 154.5;  DB 2;  Length 119;
  Best Local Similarity   32.8%;  Pred. No. 0.0019;
  Matches   38;  Conservative   12;  Mismatches   39;  Indels   27;  Gaps    3;

Qy          9 VVFDPSNPTTFSVEAIAAYTPVALIRLLNASGPLQPGHRVDIADARSIYT---------- 58
              | ||| :||           || :: :||          :|  |    ||
Db         12 VAFDPWDPTNIKAANFKDMLPVDVMTILN--------QNIDELD----YTKYSDDEINEG 59

Qy         59 -----VGAAASAARARANHNANTIRRTAMFAETDPMTWLRPTVGLKRTFNPRIIRP 109
                  :| | :    | |    :||:||||  |  || || :|||||| || ::|
Db         60 LKQLFMGTAKTMVVLRQRHLKALVRRSAMFAHNDASTWARPNIGLKRTFPPRFMQP 115


RESULT 3
WZBEB7
13.6K capsid protein - equine herpesvirus 1 (strain Ab4p)
C;Species: equine herpesvirus 1
A;Note: host Equus caballus (domestic horse)
C;Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 09-Jul-2004
C;Accession: H36797
R;Telford, E.A.R.; Watson, M.S.; McBride, K.; Davison, A.J.
submitted to GenBank, March 1992
A;Description: The DNA sequence of equine herpesvirus-1.
A;Reference number: A36805
A;Accession: H36797
A;Molecule type: DNA
A;Residues: 1-119 <TEL>
A;Cross-references: UNIPROT:P28974; GB:M86664; NID:g330791; PIDN:AAB02460.1; PID:g330817
R;Telford, E.A.R.; Watson, M.S.; McBride, K.; Davison, A.J.
Virology 189, 304-316, 1992
A;Title: The DNA sequence of equine herpesvirus-1.
A;Reference number: A41831; MUID:92295566; PMID:1318606
A;Contents: annotation; possible protein-coding frames
A;Note: neither amino acid nor nucleotide sequence is given
C;Genetics:
A;Gene: 25
C;Superfamily: human herpesvirus 1 UL35 protein; human herpesvirus 1 UL35 protein homolo
C;Keywords: capsid protein
F;19-110/Domain: human herpesvirus 1 UL35 protein homology <UL35>

  Query Match             12.7%;  Score 153.5;  DB 1;  Length 119;
  Best Local Similarity   31.1%;  Pred. No. 0.0022;
  Matches   37;  Conservative   16;  Mismatches   45;  Indels   21;  Gaps    3;

Qy          3 QPASSRVVFDPSNPTTFSVEAIAAYTPVALIRLLNASGPLQPGHRVDIADARSIYT---- 58
              |   : | |:|::|           || :| :||          :|  |    ||
Db          6 QQQQAPVAFNPADPPNIKAANFKDMLPVDVITILN--------QNIDELDYTK-YTEDEI 56

Qy         59 --------VGAAASAARARANHNANTIRRTAMFAETDPMTWLRPTVGLKRTFNPRIIRP 109
                    :| | :    | |  : :||: |||: |  || || :|||||| || ::|
Db         57 SEGLKQLFMGTARTMVSLRQRHLKSLVRRSDMFAQNDASTWARPNIGLKRTFPPRFMQP 115


RESULT 4
QQBE22
membrane antigen gp220 - human herpesvirus 4 (strain B95-8)
C;Species: human herpesvirus 4, Epstein-Barr virus
C;Date: 25-Feb-1985 #sequence_revision 25-Feb-1985 #text_change 09-Jul-2004
C;Accession: B43042; A03763
R;Bankier, A.T.; Deininger, P.L.; Farrell, P.J.; Barrell, B.G.
Mol. Biol. Med. 1, 21-45, 1983
A;Title: Sequence analysis of the 17,166 bp EcoRI fragment C of B95-8 Epstein-Barr virus
A;Reference number: A93065; MUID:85035713; PMID:6092825
A;Accession: B43042
A;Molecule type: DNA
A;Residues: 1-710 <BAN>
A;Cross-references: UNIPROT:P03200
R;Baer, R.; Bankier, A.T.; Biggin, M.D.; Deininger, P.L.; Farrell, P.J.; Gibson, T.J.; Ha
Nature 310, 207-211, 1984
A;Title: DNA sequence and expression of the B95-8 Epstein-Barr virus genome.
A;Reference number: A03794; MUID:84270667; PMID:6087149
A;Contents: annotation; protein coding region
C;Superfamily: Epstein-Barr virus membrane antigen gp220

  Query Match             11.5%;  Score 139;  DB 1;  Length 710;
  Best Local Similarity   25.5%;  Pred. No. 0.11;
  Matches   56;  Conservative   25;  Mismatches  101;  Indels   38;  Gaps    9;

Qy         37 NASGPLQPGHRVDIADARSIYTVGAAASAAR----ARANHNANTIRRTAMFAETDPMTWL 92
              | ||          ||     ||    :| :      |   || |     :|::    |
Db        378 NISGAFASNRTFDI-------TVSGLGTAPKTLIITRTATNATTTTHKVIFSKAPESTTT 430

Qy         93 RPTVGLKRTFNPRIIRPQPPNPSMSLGISGPTILPQKTQSADQSALQQPAALAFSGSSPQ 152
              ||:      :|       | :  :   :: |        :|| ::    ||    ||:||
Db        431 SPTLNTTGFADPNTTTGLPSSTHVPTNLTAPASTGPTVSTADVTS-PTPAGTT-SGASPV 488

Qy        153 HPPP--------------QTTSASVGQQQHVVSGSSGQQPQQGAQSSTVQP---TTGSPP 195
              | |               | | ||:    | || ||   |: |  |    |  :| : |
Db        489 TPSPSPWDNGTESGENITQVTPASI--STHHVSTSS-PAPRPGTTSQASGPGNSSTSTKP 545

Qy        196 AAQGVPQSTPPPTQNTPQGGKGQ-----TLSHTGQSGNAS 230
                : |||    :||   ||      |:: ||   |::
Db        546 GEVNVTKGTPPQNATSPQAPSGQKTAVPTVTSTGGKANST 585


RESULT 5
T34434
hypothetical protein K06A9.1a - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C;Accession: T34434
R;Geisel, C.; Gattung, S.
submitted to the EMBL Data Library, December 1996
A;Description: The sequence of C. elegans cosmid K06A9.
A;Reference number: Z21525
A;Accession: T34434
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-2232 <GEI>
A;Cross-references: UNIPROT:Q8IFX6; EMBL:U80846; PIDN:AAC70890.1; GSPDB:GN00028; CESP:K06
A;Experimental source: strain Bristol N2; clone K06A9
C;Genetics:
A;Gene: CESP:K06A9.1a
A;Map position: X
A;Introns: 38/1; 75/3; 103/3; 132/2; 158/2; 222/1; 1088/1; 1367/1; 2039/1; 2049/1; 2075/3

  Query Match             10.9%;  Score 131.5;  DB 2;  Length 2232;
  Best Local Similarity   26.0%;  Pred. No. 1.1;
  Matches   69;  Conservative   29;  Mismatches  130;  Indels   37;  Gaps   10;

Qy          2 TQPASSRVVFDP-SNPTTFSVEAIAAYTPVALIRLLNASGPLQPGHRVDIADARSIYTVG 60
              |  || :   |  : :|  ||:    :| :    | :: |  |      |   :   : |
Db        859 TSSGSSMITQSPYPSQSTSPVESSTTPSPGSPGTTLTSTSP-SPSQSTTIGSTQGSTSPG 917

Qy         61 AAASAARARANHNANTIRRTAMFAETDPMT-------WLRPTVG-LKRTFNPRIIRPQPP 112
               : ::    :  :  |   |     | | |           |||  : : :|     |
Db        918 ISTTSEEMTSQGSTQTPGSTGS-TVTQPSTVSDSTSSGSTVTVGSTEGSSSPIPSTSQNT 976

Qy        113 NPSMSLGISGPTILPQKTQSADQSALQQPAALAFSGS--------SPQHPPPQTTSASVG 164
              |||  | | |  |  || :||           | : |||         ||   |  |  :| |
Db        977 NPSTSSGSSMSTQTPQSSQSTSPVESSTSGATSSSGSPGTTLTSISPSPSPSSTIGSSQG 1036

Qy        165 QQQHVVSG-SSGQQPQQGAQSSTV-QPTTGSPPAAQGV---------------PQSTPPP 207
                 |||  | |       |:  ||| :|:| |  |: |                  ::| |
Db       1037 STSPVVSTISQGSTETPGSTGSTVTKPSTVSGSASSGSTATMGSTEASSTSGGSSTSPNP 1096

Qy        208 TQNTPQGGKGQTLSHTGQSGNASRS 232
```

```
              :|:|     | | |  | ||    |
Db       1097 SQSTSPSTSGAT-SSPGSSGTTLTS 1120


RESULT 6
PIHUSD
salivary proline-rich glycoprotein precursor PRB4 (large allele) [validated] - human
N;Contains: basic proline-rich protein IB-5; proline-rich peptide P-D
C;Species: Homo sapiens (man)
C;Date: 19-Feb-1984 #sequence_revision 12-Apr-1996 #text_change 09-Jul-2004
C;Accession: S03176; S03175; S10890; D25372; E38355; A03295; A61294; S62891
R;Lyons, K.M.; Stein, J.H.; Smithies, O.
Genetics 120, 267-278, 1988
A;Title: Length polymorphisms in human proline-rich protein genes generated by intrageni
A;Reference number: S02127; MUID:89121440; PMID:2851479
A;Accession: S03176
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 35-310 <LY1>
A;Cross-references: UNIPROT:P10163; UNIPROT:P10162; EMBL:X07715
A;Note: large allele
A;Accession: S03175
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 35-36,'E',38-112,155-310 <LY2>
A;Cross-references: EMBL:X07704
A;Note: medium allele
A;Accession: S10890
A;Status: preliminary; translation not shown
A;Molecule type: DNA
A;Residues: 1-38,60-112,'T',114-115,'P',117-121,185-271,'A',273-310 <LY3>
A;Cross-references: EMBL:X07882; NID:g35647; PIDN:CAA30729.1; PID:g296670
R;Maeda, N.; Kim, H.S.; Azen, E.A.; Smithies, O.
J. Biol. Chem. 260, 11123-11130, 1985
A;Title: Differential RNA splicing and post-translational cleavages in the human salivar
A;Reference number: A92492; MUID:85289325; PMID:2993301
A;Accession: D25372
A;Molecule type: mRNA
A;Residues: 1-36,'E',38-112,'T',114-115,'P',117-121,185-271,'A',273-310 <MAE>
R;Kauffman, D.L.; Bennick, A.; Blum, M.; Keller, P.J.
Biochemistry 30, 3351-3356, 1991
A;Title: Basic proline-rich proteins from human parotid saliva: relationships of the cov
A;Reference number: A38355; MUID:91190884; PMID:1849422
A;Accession: E38355
A;Molecule type: protein
A;Residues: 241-254,'KN',257-310 <KAU>
R;Saitoh, E.; Isemura, S.; Sanada, K.
J. Biochem. 93, 495-502, 1983
A;Title: Complete amino acid sequence of a basic proline-rich peptide, P-D, from human p
A;Reference number: A03295; MUID:83186122; PMID:6841349
A;Accession: A03295
A;Molecule type: protein
A;Residues: 241-310 <SAI>
R;Shimomura, H.; Kanai, Y.; Sanada, K.
J. Biochem. 93, 857-863, 1983
A;Title: Amino acid sequences of glycopeptides obtained from basic proline-rich glycoprc
A;Reference number: A61294; MUID:83265671; PMID:6874667
A;Accession: A61294
A;Molecule type: protein
A;Residues: 54-57,'E',59-73,'R';82-101 <SHI>
R;Charlton, A.J.; Baxter, N.J.; Lilley, T.H.; Haslam, E.; McDonald, C.J.; Williamson, M.
FEBS Lett. 382, 289-292, 1996
A;Title: Tannin interactions with a full-length human salivary proline-rich protein disp
A;Reference number: S62891; MUID:96184506; PMID:8605987
A;Accession: S62891
A;Molecule type: protein
A;Residues: 241-252 <CHA>
A;Note: amino end of peptide designated basic proline-rich protein IB-5
A;Note: it is unclear from the peptide sequence whether this is a product of the PRB2 (P
C;Genetics:
A;Gene: GDB:PRB4
A;Cross-references: GDB:119514; OMIM:180990
A;Map position: 12p13.2-12p13.2
A;Introns: 22/1; 34/1
A;Note: the list of introns may be incomplete
C;Superfamily: proline-rich protein
C;Keywords: glycoprotein; saliva; tandem repeat
F;1-16/Domain: signal sequence #status predicted <SIG>
F;241-310/Product: proline-rich peptide P-D #status experimental <MAT>
F;66,87,171/Binding site: carbohydrate (Asn) (covalent) #status experimental
F;108,150,192,213,234/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match            10.6%;  Score 128;  DB 1;  Length 310;
  Best Local Similarity   25.7%;  Pred. No. 0.23;
  Matches   56;  Conservative   16;  Mismatches   98;  Indels   48;  Gaps    8;

Qy         30 VALIRLLNASGPLQPGHRVDIADARSIYTVGAAASAARARANHNANTIRRTAMFAETDPM 89
              |||: | :|    :     |::   |:: :       | :  :
Db          8 VALLALSSAESSSE-----DVSQEESLFLISGKPQGRRPQGGNQPQ-------------- 48

Qy         90 TWLRPTVGLKRTFNPRIIRPQPPNPSMSLGISGPTILPQKTQSADQSALQQPAALAFSGS 149
                ||          |   :|| | |       ||   | | :       |
Db         49 ---RPP--------PPPGKPQGPPPQGGNQSQGPPPPPGKPEGRPPQGGNQSQGPPPHPG 97

Qy        150 SPQHPPPQTTSASVGQQQHVVSGSSGQQPQQGAQSSTVQPTTGSP--PAAQGVPQS--TP 205
               |: |||| : | |   |     | : || | ||    || | |  |  ||  ||  ||
Db         98 KPERPPPQGGNQSQGPPPHPGKPES-RPPQGGHQSQGPPPTPGKPEGPPPQGGNQSQGTP 156

Qy        206 P----PTQNTPQGGK---------GQTLSHTGQSGNAS 230
               |    |    ||||          |:      | || |
Db        157 PPPGKPEGRPPQGGNQSQGPPPHPGKPERPPPQGGNQS 194


RESULT 7
T18535
high molecular mass nuclear antigen - chicken (fragment)
C;Species: Gallus gallus (chicken)
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T18535
R;Shimada, K.; Harata, M.; Mizuno, S.
J. Cell Sci. 110, 3031-3041, 1997
A;Title: A nuclear matrix-associated high molecular mass nuclear antigen, HMNA, of chicke
A;Reference number: Z18955; MUID:9803440; PMID:9365273
A;Accession: T18535
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1151 <SHI>
A;Cross-references: UNIPROT:O57580; EMBL:D88440; NID:d1177138; PID:d1025045; PIDN:BAA2413

  Query Match            10.6%;  Score 128;  DB 2;  Length 1151;
  Best Local Similarity   28.2%;  Pred. No. 0.92;
  Matches   57;  Conservative   22;  Mismatches   81;  Indels   42;  Gaps   11;

Qy         59 VGAA---------ASAARARANHNANTIRRTAMFAETDPMTWLRPTVGLKRTFNPRIIRP 109
              :|||          |:| :  |     ||:: : | | | | |   |:||   |       |
Db        776 MGAAATTQPSPMGAAATQVTATSAGNTMQVSPMGAATPPQT---PSVGAATT-------P 825

Qy        110 QPPNPSMSLGISGPTILPQKTQSADQSAL-QQPAALAFSGSSPQHPP-----PQTT--SA 161
              ||      :  :  |       | :   |:  ||  :| : :: | ||      ||:|   |
Db        826 QPSPMGAATTLMSPMGAATTPQPSPMGAVTTQPPPMAATNTT-QPPPMAASTPQSTPMGA 884

Qy        162 SVGQQQHVVSGSSGQQPQQGAQSSTVQPT-TGSP---------PAAQGVPQ---STPPPT 208
              :   |   :  :: | |  || :    ||  |||         | ||  ||    :|||
Db        885 ATTTQSPPMGATTTQSPPMGASTPQAPPTVAGSPTPPPPIPPSPTAQTSPQPMSKSPPPD 944

Qy        209 QNTPQGGKGQTLSHTGQSGNAS 230
                 || |      |||
Db        945 PPKAPSAAAQT-SPAAHVANAS 965


RESULT 8
E70766
hypothetical protein Rv2082 - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C;Accession: E70766
```

```
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Sqares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: E70766
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-721 <COL>
A;Cross-references: UNIPROT:Q10690; GB:Z73966; GB:AL123456; NID:g3261577; PIDN:CAA98194.
A;Experimental source: strain H37Rv
C;Genetics:
A;Gene: Rv2082

  Query Match             10.5%;  Score 127;  DB 2;  Length 721;
  Best Local Similarity   23.4%;  Pred. No. 0.65;
  Matches   60;  Conservative   21;  Mismatches   83;  Indels   92;  Gaps   10;

Qy         34 RLLNASGPLQPGHRVDIADARSIYTVGAAASAARARANH-----NANTIRRT-AMFAET- 86
              ::|   ||::           :  | |    : | |:|     :| |  |  :| ||
Db        117 QILAGKGPIE----------AKVAAVNAVIEQSNAMADHVGATAMSNIIDATQRVFDETI 166

Qy         87 --DPMTWLRP-TVGLKRTFNPRIIRPQ-----------------PPNPSMSLGISGPTI 125
                |  ||||   | |     || :  :                |  || |   |||
Db        167 GGDAHTWLRDHGVSLDTPARPRPVTAEDMTSMTANSPAGSPFGAAPSAPSHSTTTSGPPT 226

Qy        126 LPQKTQSADQSALQQPAALAFSGSS-----------------PQHPPPQTTS------ 160
               |  |     :    |  |: | :|                 |  ||| | |
Db        227 APTPTSPFGTA----PMVLSSSSTSSGPPTAPTPTSPFGTAPMPPGPPPPGTVSPPLPPS 282

Qy        161 ----------ASVGQQQHVVSGSSGQQPQQGAQSSTVQPTTGSPPAA----------- 197
                       : |     : ::   ||   || |   |||:| ||
Db        283 APAVGVGGPSVPAAGMPPAAAAATAPLSPQSLGQSFTTGMTTGTPAAAGAQALSAGALHA 342

Qy        198 --QGVPQSTPPPTQNT 211
                : :|   ||||  |
Db        343 ATEPLPPPAPPPTTPT 358


RESULT 9
T00366
hypothetical protein KIAA0669 - human
C;Species: Homo sapiens (man)
C;Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 09-Jul-2004
C;Accession: T00366
R;Ishikawa, K.; Nagase, T.; Suyama, M.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura, N.
DNA Res. 5, 169-176, 1998
A;Title: Prediction of the coding sequences of unidentified human genes. X. The complete
A;Reference number: Z14142; MUID:98403880; PMID:9734811
A;Accession: T00366
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-780 <ISH>
A;Cross-references: UNIPROT:O75157; EMBL:AB014569; NID:g3327151; PIDN:BAA31644.1; PID:g3
A;Experimental source: brain; clone HK02346
C;Genetics:
A;Note: KIAA0669

  Query Match             10.5%;  Score 127;  DB 2;  Length 780;
  Best Local Similarity   26.3%;  Pred. No. 0.71;
  Matches   68;  Conservative   17;  Mismatches  106;  Indels   68;  Gaps   12;

Qy          1 MTQPASSRVVFDPSNPTTFSVEAIAAYTPVALIRLLNASGPLQP------GHRVDIADAR 54
              |:||        |: | :|||       |     :  :| || |      | :  :| |:
Db        258 MSQPT-------PAQPQSFSVGQPQPPPPPVGGAVAQSSAPLPPFPGAATGPQPMMAAAQ 310

Qy         55 SIYTVGAAASAARARANHNANTIRRTAMFAETDPMTWLRPTVGLKRTFNP-RIIRPQPPN 113
                ||          :   |: : ||     |     | ||      | :   |||
Db        311 PSQPQGAGPGGQTLPPTN--VTLAQPAMSLPPQP----GPAVGAPAAQQPQQFAYPQPQI 364

Qy        114 PSMSLGISGPTILPQKTQSADQSALQQPAALAFSGSSPQHP--------PPQTTSASVGQ 165
                |   |       ||   | : ||   |          ||  |        |    :  ||
Db        365 PPGHL-------LP--VQPSGQSEYLQQHVAGLQPPSPAQPSSTGAAASPATAATLPVGT 415

Qy        166 QQHVVS------GSSGQQ---------PQQGAQSSTVQPTTGSP------------PAAQ 198
               |:  |      |:| |           | || | :  |||   |             ||
Db        416 GQNASSVGAQLMGASSQPSEAMAPRTGPAQGGQVAPCQPTGVPPATVGGVVQPCLGPAGA 475

Qy        199 GVPQSTPPPTQNTPQGGKG 217
               | ||| |||      || |
Db        476 GQPQSVPPP----QMGGSG 490


RESULT 10
I49759
hepatocyte growth factor-regulated tyrosine kinase substrate Hrs [imported] - mouse
C;Species: Mus musculus (house mouse)
C;Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 09-Jul-2004
C;Accession: I49759
R;Komada, M.; Kitamura, N.
Mol. Cell. Biol. 15, 6213-6221, 1995
A;Title: Growth factor-induced tyrosine phosphorylation of Hrs, a novel 115-kilodalton pr
A;Reference number: A57554; MUID:96026002; PMID:7565774
A;Accession: I49759
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-775 <RES>
A;Cross-references: UNIPROT:Q61691; GB:D50050; NID:g1089780; PIDN:BAA08768.1; PID:g108978

  Query Match             10.4%;  Score 126;  DB 2;  Length 775;
  Best Local Similarity   25.9%;  Pred. No. 0.81;
  Matches   63;  Conservative   31;  Mismatches   83;  Indels   66;  Gaps   14;

Qy          4 PASSRVVFDPSNPTTFSVEAIAAYTP-----------VALIRLLNASGPL--QPGHRVDI 50
              | :  |:: || ||:|     | ::|           | : :   |:||    ||   |
Db        572 PTAGGVLYQPSGPTSFP----ATFSPAGSVEGSPMHGVYMSQPAPATGPYPSMPGTTADP 627

Qy         51 ADARS-IYTVGAAASAARARANHNANTIRRTAMFAETDPMTWLRPTVGLKRTFNPRIIRP 109
              :   : :|  ||  : |  :|      |     : :  |       || | :
Db        628 SMVSAYMYPTGAPGAQAAPQAQAGPTTSPAYSSYQPT-------PTPGYQ---------- 670

Qy        110 QPPNPSMSLGISGPTILPQKTQSADQSAL----QQPAALAFSGSSPQH----PPPQTTSA 161
                     |:    |  ||  :|    | :     || :: :   : |:     | |   |
Db        671 -------SVASQAPQSLPAISQPPQTSNIGYMGSQPMSMGYQPYNMQNLMTALPGQ--DA 721

Qy        162 SVGQQQHVVSGSSGQQP--QQGAQSSTVQPTTGSPPAAQGVPQSTPPPTQNTP-QGGKGQ 218
              |:  ||  :   ||||  || |      |:||  |   |  | : | |||  | || : |
Db        722 SLPAQQPYI---PGQQPLYQQMA------PSTG--PPQQQPPVAQPAPTQGPPAQGSEAQ 770

Qy        219 TLS 221
               :|
Db        771 LIS 773


RESULT 11
T29265
hypothetical protein C01G8.7 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T29265
R;Du, Z.; Gattung, S.
submitted to the EMBL Data Library, November 1996
A;Description: The sequence of C. elegans cosmid C01G8.
A;Reference number: Z20597
A;Accession: T29265
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1357 <DUZ>
A;Cross-references: EMBL:U80439; PIDN:AAB37645.1; GSPDB:GN00019; CESP:C01G8.7
A;Experimental source: strain Bristol N2; clone C01G8
C;Genetics:
A;Gene: CESP:C01G8.7
A;Map position: 1
A;Introns: 89/1; 488/1; 701/3; 1056/2; 1159/3; 1197/1; 1312/3
```

```
  Query Match             10.4%;  Score 125;  DB 2;  Length 1357;
  Best Local Similarity   23.5%;  Pred. No. 1.7;
  Matches   64;  Conservative   22;  Mismatches  122;  Indels   64;  Gaps   10;

Qy          2 TQPASSRVVFDPSNPTTFSVEAIAAYTPVALIRLLNASGPLQPGHRVDIA--------DA 53
              ||||:      |  | : |        | ::      :|  | | |   |        |:
Db        238 TQPATPSASNQPPTPASASASTTLHVEPASV------AGGSQAGSRAPSAGPGPSATPDS 291

Qy         54 RSIYTVGAAASAARARANHNANTIRRT--------AMFAETDPMTWLRPTVGLKRTFNPR 105
               :| :  ||| |    ::  :     |        |      | | ::|    :   |
Db        292 STIPSASAAADAQTTTSSETSTPGEPTSSTPSDTPAPSTSVPPPTSVQPPQPQQPPQGP- 350

Qy        106 IIRPQPPNPSMSLGISGPTILPQKTQSADQSALQQPAALA-----------------FS 147
                | || |        |   |         |:: ||  |                  |
Db        351 ---PGPPGPQQH-----PGPYPGYPGYGPPGAMRPPAGFAPPPGAPYGYPPGAPPPAGFH 402

Qy        148 GSSPQHPPPQTTSASVGQQQHVVSGSSGQQPQQGAQSSTVQPTTGSP--PAAQGVPQS-- 203
              | |||  ||       |||        || ||       :|    |  |   | ||:
Db        403 PSHPQH--PQHAQYLAWQQQRYHQQQQHQQQQQQGAPGGPRPPYPYPGGPVPPGPPQNRM 460

Qy        204 -TPPPTQNTPQGGKGQTLSHTGQSGNASRSRR 234
               ||| |  |       | :| :|:  :  |
Db        461 PPPPPAQGAP--------SPSGAAGSNGKQPR 484


RESULT 12
QQBE21
membrane antigen gp350 - human herpesvirus 4 (strain B95-8)
C;Species: human herpesvirus 4, Epstein-Barr virus
C;Date: 25-Feb-1985 #sequence_revision 25-Feb-1985 #text_change 09-Jul-2004
C;Accession: A43042; S33008; S33009; A03762
R;Bankier, A.T.; Deininger, P.L.; Farrell, P.J.; Barrell, B.G.
Mol. Biol. Med. 1, 21-45, 1983
A;Title: Sequence analysis of the 17,166 bp EcoRI fragment C of B95-8 Epstein-Barr virus
A;Reference number: A93065; MUID:85035713; PMID:6092825
A;Accession: A43042
A;Molecule type: DNA
A;Residues: 1-907 <BAN>
A;Cross-references: UNIPROT:P03200; GB:V01555; GB:J02070; GB:K01729; GB:K01730; GB:V0155
R;Farrell, P.J.
submitted to the EMBL Data Library, March 1988
A;Reference number: S32973
A;Accession: S33008
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-907 <FAR>
A;Cross-references: EMBL:V01555; NID:g59074; PIDN:CAA24855.1; PID:g1334869
A;Accession: S33009
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-907 <FA2>
A;Cross-references: EMBL:V01555; NID:g59074; PIDN:CAA24855.1; PID:g1334869
R;Baer, R.; Bankier, A.T.; Biggin, M.D.; Deininger, P.L.; Farrell, P.J.; Gibson, T.J.; H
Nature 310, 207-211, 1984
A;Title: DNA sequence and expression of the B95-8 Epstein-Barr virus genome.
A;Reference number: A03794; MUID:84270667; PMID:6087149
A;Contents: annotation; protein coding region
C;Superfamily: Epstein-Barr virus membrane antigen gp350

  Query Match             10.3%;  Score 124.5;  DB 1;  Length 907;
  Best Local Similarity   24.9%;  Pred. No. 1.2;
  Matches   64;  Conservative   25;  Mismatches   89;  Indels   79;  Gaps   14;

Qy          2 TQPASSRVVFDP--SNPT---TFSVEAIAAYTPVALIRLLNASGPLQPGHRVDIADARSI 56
              | | |:     |  ::||   |    |:   ||       ||:||
Db        577 TSPTSAVTTPTPNATSPTLGKTSPTSAVTTPTP-------NATGP-------------- 614

Qy         57 YTVGAAASAARARANHNANTIRRTAMFAETDPMTWLRPTVGLKRTFNPRIIRPQPPNPSM 116
               |||  :  | |  ||   |:  |:           ||          ::  || | :
Db        615 -TVGETSPQANA-TNH---TLGGTS-----------PT---------PVVTSQPKNATS 648

Qy        117 SLGISGPTILPQKTQSADQSALQQPAAL--AFSGSSPQHPP-------------PQTTSA 161
```

```
              ::      |    | |        |  |  : | :|  | |                 | | |
Db        649 AVTTGQHNITSSSTSSMSLRPSSNPETLSPSTSDNSTSHMPLLTSAHPTGGENITQVTPA 708

Qy        162 SVGQQQHVVSGSSGQQPQQGAQSSTVQP---TTGSPPAAQGVPQSTPPPTQNTPQGGKGQ 218
              |:    | || ||   |: | |    |   :| : |    | : |||     :||    ||
Db        709 SI--STHHVSTSS-PAPRPGTTSQASGPGNSSTSTKPGEVNVTKGTPPQNATSPQAPSGQ 765

Qy        219 -----TLSHTGQSGNAS 230
                  |:: ||   |::
Db        766 KTAVPTVTSTGGKANST 782


RESULT 13
T34433
hypothetical protein K06A9.1a - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C;Accession: T34433
R;Geisel, C.; Gattung, S.
submitted to the EMBL Data Library, December 1996
A;Description: The sequence of C. elegans cosmid K06A9.
A;Reference number: Z21525
A;Accession: T34433
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1032 <GEI>
A;Cross-references: UNIPROT:P91365; EMBL:U80846; PIDN:AAC70889.1; GSPDB:GN00028; CESP:K0
A;Experimental source: strain Bristol N2; clone K06A9
C;Genetics:
A;Gene: CESP:K06A9.1a
A;Map position: X
A;Introns: 38/1; 75/3; 103/3; 132/2; 158/2; 222/1; 839/1; 849/1; 875/3; 940/2; 966/2; 99

  Query Match             10.3%;  Score 124.5;  DB 2;  Length 1032;
  Best Local Similarity   23.3%;  Pred. No. 1.4;
  Matches   59;  Conservative   33;  Mismatches  108;  Indels   53;  Gaps   11;

Qy          7 SRVVFDPSNPTTFSVEAIAAYTPVALIRLLNASGPLQPGHRVDIADARSIYTVGAAASAA 66
              | |    | : :||      : ||:|            |  | :    |  | |::  :|
Db        419 STVTVAPVSSSTF-----GSSTPIA--------SSSSSGSTVTVVSGSS-STYGSSTPSA 464

Qy         67 RARANHNANTIRRTAMFAETDPMTWLRPTVGLKRTFNPRIIRPQPPNPSMSLGISGP--- 123
                : :   |:||         |     : |  |   :          | :|     :||
Db        465 SSSSAGTASTIS-----GSTGSTATIVP--GSSSSVGSSTQSASPSSPGTMSTVSGPTGS 517

Qy        124 --TILPQKTQSADQSALQQPAA--------LAFSGSSP---QHPPPQTTSASVGQQQHVV 170
                 |::|  : |   |:   |::       :  ||||        | | | |  :
Db        518 TVTVVPGSSTSPAPSSSPNPSSSPASTGSTITISGSSSIIVSTVSGSTVSGSTGTSQSTL 577

Qy        171 SGS----------SGQQPQQGAQSSTVQPTTGSPPAAQGVPQSTPP---PTQNTPQGGK 216
              : |          |   ||  :||    | |||   :|   ||  |   |: :|| |
Db        578 ASSTATPGSSSTVPSSSSPQPSSQSPA--PNTGSTTPSQTSSQSPSPSMNPSSSTPTGSS 635

Qy        217 GQTLSHTGQSGNA 229
                |::  | : ::
Db        636 QSTITPEGSTASS 648


RESULT 14
PIHUB6
salivary proline-rich phosphoprotein precursor PRB1 (large allele) [validated] - human
N;Contains: peptide IB-1; peptide P-E (peptide IB-9); peptide P-F; peptide P-H
C;Species: Homo sapiens (man)
C;Date: 04-Dec-1986 #sequence_revision 12-Apr-1996 #text_change 09-Jul-2004
C;Accession: B40750; C40750; A40750; C25372; S02128; S02127; A03293; A90502; A91974; A05
R;Azen, E.A.; Latreille, P.; Niece, R.L.
Am. J. Hum. Genet. 53, 264-278, 1993
A;Title: PRB1 gene variants coding for length and null polymorphisms among human salivary
A;Reference number: A40750; MUID:93304421; PMID:8317492
A;Accession: B40750
A;Molecule type: DNA
A;Residues: 35-392 <AZE>
A;Cross-references: UNIPROT:P04280; UNIPROT:P04281; UNIPROT:P02811; UNIPROT:Q16038; GB:S
```

A;Experimental source: subject C.J. (large allele)
A;Accession: C40750
A;Molecule type: DNA
A;Residues: 35-127,'R',129-148,'R',150-151,153-187,'K',189-272,'S',274-336,'S',338-392 <
A;Cross-references: GB:S62929
A;Experimental source: subject M.V.O. (large allele)
A;Accession: A40750
A;Molecule type: DNA
A;Residues: 35-183,245-270,'Q',272-392 <AZ3>
A;Cross-references: GB:S62928
A;Experimental source: subject C.J. (medium allele)
A;Note: authors translated the codon CAA for residue 272 as Arg
R;Maeda, N.; Kim, H.S.; Azen, E.A.; Smithies, O.
J. Biol. Chem. 260, 11123-11130, 1985
A;Title: Differential RNA splicing and post-translational cleavages in the human salivar
A;Reference number: A92492; MUID:85289325; PMID:2993301
A;Accession: C25372
A;Molecule type: mRNA
A;Residues: 1-183,245-392 <MAE>
A;Cross-references: GB:K03204; NID:g190485; PIDN:AAA60185.1; PID:g190486
A;Note: alternatively splice forms lacking portions of the repeat region were also found
R;Lyons, K.M.; Stein, J.H.; Smithies, O.
Genetics 120, 267-278, 1988
A;Title: Length polymorphisms in human proline-rich protein genes generated by intrageni
A;Reference number: S02127; MUID:89121440; PMID:2851479
A;Accession: S02128
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 35-127,250-273,'R',275-277,'R',279-336,'S',338-392 <LYO>
A;Cross-references: EMBL:X07517
A;Accession: S02127
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 35-183,245-392 <LY2>
A;Cross-references: EMBL:X07516
R;Kauffman, D.; Hofmann, T.; Bennick, A.; Keller, P.
Biochemistry 25, 2387-2392, 1986
A;Title: Basic proline-rich proteins from human parotid saliva: complete covalent struct
A;Reference number: A90502; MUID:86243355; PMID:3521730
A;Accession: A03293
A;Molecule type: protein
A;Residues: 17-38,'AP',41-51,92-148,'R',150-152 <KA2>
A;Note: among nine basic proline-rich peptides isolated from the saliva, this peptide is
A;Accession: A90502
A;Molecule type: protein
A;Residues: 275-336,'S',338-392 <KAU>
R;Saitoh, E.; Isemura, S.; Sanada, K.
J. Biochem. 94, 1991-1999, 1983
A;Title: Further fractionation of basic proline-rich peptides from human parotid saliva
A;Reference number: A91974; MUID:84161824; PMID:6671974
A;Contents: P-H
A;Accession: A91974
A;Molecule type: protein
A;Residues: 'S',338-392 <SAI>
R;Azen, E.; Lyons, K.M.; McGonigal, T.; Barrett, N.L.; Clements, L.S.; Maeda, N.; Vanin,
Proc. Natl. Acad. Sci. U.S.A. 81, 5561-5565, 1984
A;Reference number: A94005; MUID:84298176; PMID:6089212
A;Accession: A05261
A;Molecule type: DNA
A;Residues: 35-39,'P',41-84,'G',86,'R',87-154,'R',218-246;300-306,'T',308-329,'C',331-38
A;Accession: A05262
A;Molecule type: DNA
A;Residues: 'N',57-59,'A',61-69;334-336,'S',338-339,'R',341-392 <AZ5>
R;Kauffman, D.; Wong, R.; Bennick, A.; Keller, P.
Biochemistry 21, 6558-6562, 1982
A;Title: Basic proline-rich proteins from human parotid saliva: complete covalent struct
A;Reference number: A90464; MUID:83101329; PMID:6924859
A;Contents: IB-9
A;Accession: A90464
A;Molecule type: protein
A;Residues: 92-127,'R',129-148,'R',150-152 <KA3>
R;Isemura, S.; Saitoh, E.; Sanada, K.
J. Biochem. 91, 2067-2075, 1982

A;Title: Fractionation and characterization of basic proline-rich peptides of human parot
A;Reference number: A91966; MUID:83007119; PMID:7118863
A;Contents: P-E
A;Accession: A91966
A;Molecule type: protein
A;Residues: 92-127,'R',129-148,'R',150-152 <ISE>
C;Comment: This peptide contains 21-residue repeats, two of which have internal 7-residue
C;Genetics:
A;Gene: GDB:PRB1
A;Cross-references: GDB:119511; OMIM:180989
A;Map position: 12p13.2-12p13.2
A;Note: each of the tandem repeats contains a candidate splice acceptor site, and several
C;Superfamily: proline-rich protein
C;Keywords: alternative splicing; duplication; parotid gland; phosphoprotein; pyroglutami
F;1-16/Domain: signal sequence #status predicted <SIG>
F;17-51,92-152/Product: basic proline-rich peptide IB-1 #status experimental <IB1>
F;92-152/Product: basic proline-rich peptide P-E #status experimental <PPE>
F;275-392/Product: basic proline-rich peptide IB-6 #status experimental <PIB6>
F;275-335/Product: basic proline-rich peptide P-F #status experimental <PPF>
F;337-392/Product: basic proline-rich peptide P-H #status experimental <PPH>
F;17/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experiment
F;24/Binding site: phosphate (Ser) (covalent) #status experimental

Query Match 10.2%; Score 123; DB 1; Length 392;
Best Local Similarity 32.4%; Pred. No. 0.62;
Matches 44; Conservative 7; Mismatches 59; Indels 26; Gaps 6;

```
Qy        108 RPQPPNPSMSLGISGPTILPQKTQSADQSALQQPAALAFSGSSPQHPPPQ-----TTSAS 162
              :|| | |         ||   | | |    |     :|      |   || ||||     :  :
Db        218 KPQGPPPQGGNQPQGPPPPGKPQGPPQQGGNRPQGPPPPG-KPQGPPPQGDKSRSPQSP 276

Qy        163 VGQQQHVVSGSSGQQPQQGAQSSTVQPTTGSPPAAQGVPQSTPPPTQNTPQG----GKGQ 218
               |: |       |  || | |           ||   | ||  ||   | |||    || |
Db        277 PGKPQ-------GPPPQGGNQPQ-------GPPPPPGKPQGPPPQGGNKPQGPPPPGKPQ 322

Qy        219 TLSHTGQSGNASRSRR 234
                | |: |:| |
Db        323 --GPPAQGGSKSQSAR 336
```

RESULT 15
S29605
glycoprotein 350/220 - human herpesvirus 4
C;Species: human herpesvirus 4, Epstein-Barr virus
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
C;Accession: S29605
R;Klein, K.; Mueller-Lantzsch, N.
submitted to the EMBL Data Library, October 1992
A;Description: Sequences of the membrane proteins gp 350/220 and p140 of Epstein-Barr vi
A;Reference number: S29605
A;Accession: S29605
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-886 <KLE>
A;Cross-references: UNIPROT:Q07284; EMBL:X67776; NID:g59163; PIDN:CAA47986.1; PID:g59164
C;Superfamily: Epstein-Barr virus membrane antigen gp350
C;Keywords: glycoprotein

Query Match 10.2%; Score 122.5; DB 2; Length 886;
Best Local Similarity 22.4%; Pred. No. 1.6;
Matches 56; Conservative 33; Mismatches 110; Indels 51; Gaps 8;

```
Qy          2 TQPASSRVVFDPSNPTTFSVEAIAAYTPVALIRLLNASGPLQPGHRVDIADARSIYTVGA 61
              | | |:  |  |:   |    |:   || | |  |  : |
Db        542 TSPTSA--VTTPTPNATSPTPAVTTPTPNATIPTLGKTSP-------------------- 579

Qy         62 AASAARARANHNANTIRRTAMFAETDPMTWLRPTVGLKRTFNPRIIRPQPPNPSMSLGIS 121
               ::      |  : |:  |:  | |   |       |  | :  ::   | | : ::
Db        580 TSAVTTPTPNATSPTVGETSPQANTTNHT-------LGGTSSTPVVTSPPKNATSAVTTG 632

Qy        122 GPTILPQKTQSAD--QSALQQPAALAFSGSSPQHPPPQTTSASVGQQQ----------H 168
                |    | |     |:: :  : : | :|  | |  |::   | :           |
Db        633 QHNITSSSTSSMSLRPSSISETLSPSTSDNSTSHMPLLTSAHPTGGENITQVTPASTSTH 692
```

```
Qy        169 VVSGSSGQQPQQGAQSSTVQP---TTGSPPAAQGVPQSTPPPTQNTPQGGKGQ-----TL 220
Db        693 HVSTSS-PAPRPGTTSQASGPGNSSTSTKPGEVNVTKGTPPKNATSPQAPSGQKTAVPTV 751

Qy        221 SHTGQSGNAS 230
Db        752 TSTGGKANST 761
```

Search completed: September 25, 2005, 21:03:20
Job time : 76 secs

This page blank (uspto)

OM protein - protein search, using sw model

Run on: September 25, 2005, 15:48:09 ; Search time 326 Seconds
(without alignments)
278.800 Million cell updates/sec

Title: US-09-874-140-2
Perfect score: 1206
Sequence: 1 MTQPASSRVVFDPSNPTTFS.....KGQTLSHTGQSGNASRSRRV 235

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1206	100.0	235	2	AAY25337	Aay25337 VZV VP26
2	146	12.1	1367	7	ADB70311	Adb70311 C. neofor
3	133.5	11.1	496	8	ADR08679	Adr08679 Human pro
4	127.5	10.6	907	1	AAP50073	Aap50073 Epstein-B
5	127.5	10.6	907	2	AAR80144	Aar80144 EBV gp350
6	127.5	10.6	907	3	AAY68009	Aay68009 Epstein-B
7	127	10.5	634	8	ABO84900	Abo84900 Human can
8	127	10.5	724	8	ABO84902	Abo84902 Human can
9	127	10.5	780	8	ADI82549	Adi82549 Human mod
10	126.5	10.5	839	5	AAU86145	Aau86145 Human PRO
11	126.5	10.5	839	5	ABG34056	Abg34056 Human Pro
12	126.5	10.5	839	6	ADA01322	Ada01322 Human PRO
13	126.5	10.5	839	6	ADA43751	Ada43751 Human sec
14	126.5	10.5	839	6	ADA43519	Ada43519 Human sec
15	126.5	10.5	839	6	ADA01194	Ada01194 Human PRO
16	126.5	10.5	839	7	ADA01078	Ada01078 Human sec
17	126.5	10.5	839	7	ADA43635	Ada43635 Human sec
18	126.5	10.5	839	7	ADA06897	Ada06897 Human PRO
19	126.5	10.5	839	7	ADA08385	Ada08385 Novel hum
20	126.5	10.5	839	7	ADB99678	Adb99678 Human PRO
21	126.5	10.5	839	7	ADB86961	Adb86961 Human PRO
22	126.5	10.5	839	7	ADB66116	Adb66116 Human sec
23	126.5	10.5	839	7	ADB99794	Adb99794 Human PRO
24	126.5	10.5	839	7	ADB99449	Adb99449 Novel hum
25	126.5	10.5	839	7	ADB66000	Adb66000 Human sec
26	126.5	10.5	839	7	ADC23398	Adc23398 Human tra
27	126.5	10.5	839	7	ADC26091	Adc26091 Human PRO
28	126.5	10.5	839	7	ADE04918	Ade04918 Human PRO
29	126.5	10.5	839	7	ADE11224	Ade11224 Human PRO
30	126.5	10.5	839	7	ADD88155	Add88155 Human PRO
31	126.5	10.5	839	7	ADD95450	Add95450 Human sec
32	126.5	10.5	839	7	ADE06380	Ade06380 Human PRO
33	126.5	10.5	839	7	ADE38155	Ade38155 Human PRO
34	126.5	10.5	839	7	ADD88271	Add88271 Human PRO
35	126.5	10.5	839	7	ADD90852	Add90852 Human sec
36	126.5	10.5	839	7	ADF99407	Adf99407 Human sec
37	126.5	10.5	839	7	ADG06500	Adg06500 Human PRO
38	126.5	10.5	839	7	ADG05451	Adg05451 Human PRO
39	126.5	10.5	839	7	ADG82452	Adg82452 Human PRO
40	126.5	10.5	839	7	ADJ37317	Adj37317 Human tum
41	126.5	10.5	839	8	ADE51705	Ade51705 Human sec
42	126.5	10.5	839	8	ADE51821	Ade51821 Human sec
43	126.5	10.5	839	8	ADE37679	Ade37679 Human sec
44	126.5	10.5	839	8	ADE37563	Ade37563 Human sec
45	126.5	10.5	839	8	ADD95334	Add95334 Human sec

ALIGNMENTS

RESULT 1
AAY25337
ID AAY25337 standard; protein; 235 AA.
XX
AC AAY25337;
XX
DT 03-SEP-1999 (first entry)
XX
DE VZV VP26 protein.
XX
KW VP26; immunoreactive protein; immunoassay; detection; anti-VZV; antibody.
XX
OS Human herpesvirus 3.
XX
PN DE19757765-A1.
XX
PD 24-JUN-1999.
XX
PF 23-DEC-1997; 97DE-01057765.
XX
PR 23-DEC-1997; 97DE-01057765.
XX
PA (DADE-) DADE BEHRING MARBURG GMBH.
XX
PI Eickmann M, Gicklhorn D, Radsak K, Hauser H, Giesendorf B;
XX
DR WPI; 1999-358931/31.
DR N-PSDB; AAX78759.
XX
PT Varicella zoster virus VP26 peptide - and corresponding nucleic acid,
PT useful for diagnosis of VZV infections.
XX
PS Claim 3; Page 4; 12pp; German.
XX
CC This invention describes a novel immunoreactive peptide comprising amino
CC acids 12-235 of varicella zoster virus (VZV) protein VP26, and a nucleic
CC acid encoding the peptide. Also claimed is a nucleic acid corresponding
CC to a defined DNA sequence of 705 bp given in the specification, coding
CC for amino acids 1-235 of VZV VP26. The peptide can be used in
CC immunoassays for detecting anti-VZV antibodies. The nucleic acid can be
CC used in hybridisation assays for detecting VZV
XX
SQ Sequence 235 AA;

Query Match 100.0%; Score 1206; DB 2; Length 235;
Best Local Similarity 100.0%; Pred. No. 6e-88;
Matches 235; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MTQPASSRVVFDPSNPTTFSVEAIAAYTPVALIRLLNASGPLQPGHRVDIADARSIYTVG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MTQPASSRVVFDPSNPTTFSVEAIAAYTPVALIRLLNASGPLQPGHRVDIADARSIYTVG 60

Qy         61 AAASAARARANHNANTIRRTAMFAETDPMTWLRPTVGLKRTFNPRIIRPQPPNPSMSLGI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AAASAARARANHNANTIRRTAMFAETDPMTWLRPTVGLKRTFNPRIIRPQPPNPSMSLGI 120

Qy        121 SGPTILPQKTQSADQSALQQPAALAFSGSSPQHPPPQTTSASVGQQQHVVSGSSGQQPQQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SGPTILPQKTQSADQSALQQPAALAFSGSSPQHPPPQTTSASVGQQQHVVSGSSGQQPQQ 180

Qy        181 GAQSSTVQPTTGSPPAAQGVPQSTPPPTQNTPQGGKGQTLSHTGQSGNASRSRRV 235
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GAQSSTVQPTTGSPPAAQGVPQSTPPPTQNTPQGGKGQTLSHTGQSGNASRSRRV 235
```

RESULT 2
ADB70311
ID ADB70311 standard; protein; 1367 AA.
XX
AC ADB70311;
XX
DT 04-DEC-2003 (first entry)
XX
DE C. neoformans amino acid sequence SEQ ID NO:3355.
XX
KW fungicide; gene therapy; infection.
XX
OS Cryptococcus neoformans.
XX
PN WO2003052076-A2.
XX
PD 26-JUN-2003.
XX
PF 17-DEC-2002; 2002WO-US040225.
XX
PR 17-DEC-2001; 2001US-0341261P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Zamudio C, Eroshkin AM;
XX
DR WPI; 2003-533017/50.
DR N-PSDB; ADB69228.
XX
PT New nucleic acid, useful for preparing a composition for treating an
PT infection caused by Cryptococcus neoformans.
XX
PS Claim 9; SEQ ID NO 3355; 136pp; English.
XX
CC The invention relates to a novel purified or isolated Cryptococcus
CC neoformans nucleic acid molecule comprising a sequence encoding a
CC polypeptide comprising a sequence not given in the specification. A
CC polynucleotide of the invention has fungicide activity, and may have a
CC use in gene therapy. The nucleic acid is useful for preparing a
CC composition for treating an infection caused by Cryptococcus neoformans.
CC The present sequence represents a C. neoformans sequence of the
CC invention. Note: The sequence data for this patent is not represented in
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pat_sequences.
XX
SQ Sequence 1367 AA;

```
  Query Match             12.1%;  Score 146;  DB 7;  Length 1367;
  Best Local Similarity   25.5%;  Pred. No. 0.011;
  Matches   70;  Conservative   25;  Mismatches   68;  Indels  112;  Gaps   14;

Qy          5 ASSRVVFDPSNPTTFSVEAIAAY-TPVALIRLLNASGPLQPGHRVDIADARSIYTVGAAA 63
              |||:  :    |||::     ||  | |      |||| ||               |:
Db        944 ASSQPAYALQQPTTYAASTAPAYQAPPA------ASGPYQP---------------AST 981

Qy         64 SAARARANHNANTIRRTAMFAETD---PMTWLRPTVGLKRTFNPRIIRPQPPNP------ 114
```

```
              | | |         : |  : :::   | |  :|: |    : |   ||  | |
Db        982 STAYAPP-------QPTPSYGDSNPYAPATVYQPSSG----YPPNGYRPSDPQPQGYGAP 1030

Qy        115 ----------------SMSLGISG----PTILPQKTQSADQ-----SALQQP------- 141
                              |   ||||    ||  |:: ||| :     ||:  |
Db       1031 QPSFGQTQAMPPPPRVGSQQRGISGWNDAPTFAPKRPQSAAKDVKKASAILSPFPNSADP 1090

Qy        142 -----AALAFSGSSPQ----------HPPPQTT----------SASVGQQQHVVSGSSG 175
                  | |  :|::|           |||:                :| |||     :
Db       1091 LAAAGAGLNTAGAAPPPGRSPQPGVIPPPPKNARPPSIAAKVQPPPTVQQQQQFHQHAQH 1150

Qy        176 QQPQQGAQSSTVQPTTGSPPAAQGVPQST---PPP 207
              || |:  | :       |||| | | |    |||
Db       1151 QQQQRQQQLT-------SPPAAAGPPPSAFSRPPP 1178
```

RESULT 3
ADR08679
ID ADR08679 standard; protein; 496 AA.
XX
AC ADR08679;
XX
DT 04-NOV-2004 (first entry)
XX
DE Human protein useful for treating neurological disease Seq 2185.
XX
KW human; oligo-capping method; diagnostic marker; gene therapy;
KW osteoporosis; neurological disease; Alzheimer's disease;
KW Parkinson's disease; dementia; short memory; cancer;
KW sense or motor function; emotional reaction; fear response; panic;
KW osteopathic; neuroprotective; nootropic; antiparkinsonian; cytostatic;
KW tranquiliser.
XX
OS Homo sapiens.
XX
PN EP1447413-A2.
XX
PD 18-AUG-2004.
XX
PF 12-FEB-2004; 2004EP-00003145.
XX
PR 14-FEB-2003; 2003JP-00102207.
PR 09-MAY-2003; 2003JP-00131452.
XX
PA (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
PI Isogai T, Yamamoto J, Nishikawa T, Isono Y, Sugiyama T, Otsuki T;
PI Wakamatsu A, Ishii S, Nagai K, Irie R;
XX
DR WPI; 2004-583265/57.
DR N-PSDB; ADR06723.
XX
PT New 1995 cDNA, useful for treating osteoporosis, neurological diseases,
PT Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.
XX
PS Claim 1; SEQ ID NO 2185; 2686pp; English.
XX
CC This invention relates to novel, isolated full length human cDNA
CC molecules and the encoded proteins thereof. Specifically, it refers to
CC cDNA clones obtained by an oligo-capping method, where none of these
CC clones are identical to any known human mRNAs. The present invention
CC describes an immunoassay to identify agonists and antagonists, as well as
CC antibodies, antisense molecules and siRNAs that can all be used to bind
CC to and modulate expression of the cDNA molecules. As such, these
CC molecules are useful for diagnostic markers or therapeutic targets for
CC the various diseases or morbid states. In particular, they are useful in
CC gene therapy for treating osteoporosis, neurological disease, Alzheimer's
CC disease, Parkinson's disease, dementia, short memory and various cancers,
CC as well as for maintaining equilibrium of sense or motor function, and
CC for treating emotional reaction, fear response and panic. Accordingly,
CC they exhibit osteopathic, neuroprotective, nootropic, antiparkinsonian,
CC cytostatic and tranquiliser activities. This polypeptide is a protein
CC encoded by a full length human cDNA sequence of the invention. NOTE: This

```
CC   sequence is not given in the sequence listing of the specification but
CC   can be obtained on CD-ROM from the European Patent Office, Vienna Sub-
CC   office.
XX
SQ   Sequence 496 AA;

  Query Match             11.1%;  Score 133.5;  DB 8;  Length 496;
  Best Local Similarity   26.1%;  Pred. No. 0.033;
  Matches   55;  Conservative   19;  Mismatches   58;  Indels   79;  Gaps    9;

Qy         69 RANHNANTIRRTAMFAETDPMTWLRPTVGLKRTFNPRIIRPQPPNPSMS----LGISGP- 123
              :| |  |   ||   : |:| |                   :||:| ||    || : |
Db        174 KAKHGNNQPARTGTLSRTNPPT-----------------QKPPSPPMSGRGTLGRNTPY 215

Qy        124 -TILPQKTQSADQSALQQPAALAFSGSSPQHPPPQTTSASVGQQQHVVSGSSG------- 175
               |: | |  :     :  || |   ||  || | :|  ||: |:    |||||
Db        216 KTLEPVKPPTVPNDYMTSPARL---GS--QHSPGRT--ASLNQRPRTHSGSSGGSGSREN 268

Qy        176 --------------QQPQQGAQSSTVQPTTGSPPAAQ-----------GVPQSTP---- 205
                           |  | :: :| | :|:|||                 |  |
Db        269 SGSSSIGIPIAVPTPSPPTIGPENISVPPPSGAPPAPPLAPLLPVSTVIAAPGSAPGSQY 328

Qy        206 -----------PPTQNTPQGGKGQTLSHTGQ 225
                         | :|  ||  :| | | |
Db        329 GTMTRQISRHNSTTSSTSSGGYRRTPSVTAQ 359


RESULT 4
AAP50073
ID   AAP50073 standard; protein; 907 AA.
XX
AC   AAP50073;
XX
DT   24-OCT-2003  (revised)
DT   25-MAR-2003  (revised)
DT   17-OCT-1991  (first entry)
XX
DE   Epstein-Barr virus (EBV) outer surface protein.
XX
KW   Epstein-Barr virus; antigen; vaccine.
XX
OS   Human herpesvirus 4.
XX
PN   EP151079-A.
XX
PD   07-AUG-1985.
XX
PF   28-JAN-1985;   85EP-00400141.
XX
PR   30-JAN-1984;   84US-00575352.
PR   23-JUL-1984;   84US-00633558.
XX
PA   (UYCH-) UNIV CHICAGO.
XX
PI   Kleff E,  Tanner J,  Hummel M,  Belsel C;
XX
DR   WPI; 1985-191978/32.
DR   N-PSDB; AAN50114.
XX
PT   New fragment of Epstein-Barr Virus DNA - useful in vector to express
PT   polypeptide for use in prepn. of vaccine against the virus and for use in
PT   diagnosis.
XX
PS   Claim 2; Page 23-25; 26pp; English.
XX
CC   The sequence encodes an outer surface viral protein of EBV, used to
CC   generate antibodies reacting with the surface proteins of EBV-infected
CC   cells, and in the preparation of a vaccine against EBV. (Updated on 25-
CC   MAR-2003 to correct PA field.) (Updated on 24-OCT-2003 to standardise OS
CC   field)
XX
SQ   Sequence 907 AA;

  Query Match             10.6%;  Score 127.5;  DB 1;  Length 907;
  Best Local Similarity   24.9%;  Pred. No. 0.2;
  Matches   64;  Conservative   26;  Mismatches   88;  Indels   79;  Gaps   14;

Qy          2 TQPASSRVVFDP--SNPT---TFSVEAIAAYTPVALIRLLNASGPLQPGHRVDIADARSI 56
              | | |:     |  ::||   |    |:   ||       ||:||              .
Db        577 TSPTSAVTTPTPNATSPTLGKTSPTSAVTTPTP-------NATGP-------------- 614

Qy         57 YTVGAAASAARARANHNANTIRRTAMFAETDPMTWLRPTVGLKRTFNPRIIRPQPPNPSM 116
                |||  :  | |  ||   |:  |:            ||          ::  || | :
Db        615 -TVGETSPQANA-TNH---TLGGTS-----------PT---------PVVTSQPKNATS 648

Qy        117 SLGISGPTILPQKTQSADQSALQQPAAL--AFSGSSPQHPP-------------PQTTSA 161
              ::      |   | |       |  |  : | :|  | |               | | |
Db        649 AVTTGQHNITSSSTSSMSLRPSSNPETLSPSTSDNSTSHMPLLTSAHPTGGENITQVTPA 708

Qy        162 SVGQQQHVVSGSSGQQPQQGAQSSTVQP---TTGSPPAAQGVPQSTPPPTQNTPQGGKGQ 218
              |:    | || ||  :|: |  |    |   :| : |    | : |||    :||   ||
Db        709 SI--STHHVSTSS-PEPRPGTTSQASGPGNSSTSTKPGEVNVTKGTPPQNATSPQAPSGQ 765

Qy        219 -----TLSHTGQSGNAS 230
                  |:: ||   |::
Db        766 KTAVPTVTSTGGKANST 782


RESULT 5
AAR80144
ID   AAR80144 standard; protein; 907 AA.
XX
AC   AAR80144;
XX
DT   16-OCT-2003  (revised)
DT   18-JAN-1996  (first entry)
XX
DE   EBV gp350/220.
XX
KW   EBV; gp350; gp220; gp350/gp220; non-splicing variant; vaccine.
XX
OS   Human herpesvirus 4.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1. .18
FT                   /label= Sig_peptide
FT   Region          861. .881
FT                   /note= "transmembrane region"
XX
PN   WO9528488-A1.
XX
PD   26-OCT-1995.
XX
PF   13-APR-1995;   95WO-US004611.
XX
PR   18-APR-1994;   94US-00229291.
XX
PA   (AVIR-) AVIRON.
XX
PI   Spaete R,  Jackman WT;
XX
DR   WPI; 1995-373802/48.
DR   N-PSDB; AAT04821.
XX
PT   New DNA encoding a homogeneous gp350 protein - can be used for preventing
PT   and treating Epstein-Barr virus-related diseases or conditions.
XX
PS   Disclosure; Fig 1; 61pp; English.
XX
CC   The donor and acceptor splice sites of the EBV gene encoding gp350/ 220
CC   are mutated by replacement of native nucleotides by non-native
CC   nucleotides, without altering the encoded amino acid sequence, resulting
CC   in elimination of gp220 prodn. Recombinant homogeneous gp350, useful in
CC   vaccines, is expressed in mammalian or insect cell hosts. (Updated on 16-
CC   OCT-2003 to standardise OS field)
XX
```

```
SQ   Sequence 907 AA;

  Query Match             10.6%;  Score 127.5;  DB 2;  Length 907;
  Best Local Similarity   24.9%;  Pred. No. 0.2;
  Matches   64;  Conservative   26;  Mismatches   88;  Indels   79;  Gaps   14;

Qy          2 TQPASSRVVFDP--SNPT---TFSVEAIAAYTPVALIRLLNASGPLQPGHRVDIADARSI 56
              | | |:     |  ::||   |    |:   ||       ||:||
Db        577 TSPTSAVTTPTPNATSPTLGKTSPTSAVTTPTP-------NATGP-------------- 614

Qy         57 YTVGAAASAARARANHNANTIRRTAMFAETDPMTWLRPTVGLKRTFNPRIIRPQPPNPSM 116
               |||  :  | |  ||   |:  |:            ||          ::  || | :
Db        615 -TVGETSPQANA-TNH---TLGGTS-----------PT---------PVVTSQPKNATS 648

Qy        117 SLGISGPTILPQKTQSADQSALQQPAAL--AFSGSSPQHPP------------PQTTSA 161
              ::      |    | |         |  |  : | :|  | |              | | |
Db        649 AVTTGQHNITSSSTSSMSLRPSSNPETLSPSTSDNSTSHMPLLTSAHPTGGENITQVTPA 708

Qy        162 SVGQQQHVVSGSSGQQPQQGAQSSTVQP---TTGSPPAAQGVPQSTPPPTQNTPQGGKGQ 218
              |:    | || ||  :|: |  |     |  :| : |    | : |||    :||   ||
Db        709 SI--STHHVSTSS-PEPRPGTTSQASGPGNSSTSTKPGEVNVTKGTPPQNATSPQAPSGQ 765

Qy        219 -----TLSHTGQSGNAS 230
                  |:: ||   |::
Db        766 KTAVPTVTSTGGKANST 782


RESULT 6
AAY68009
ID   AAY68009 standard; protein; 907 AA.
XX
AC   AAY68009;
XX
DT   06-AUG-2003  (revised)
DT   11-APR-2000  (first entry)
XX
DE   Epstein-Barr virus major outer envelope glycoprotein Gp350.
XX
KW   Epstein-Barr virus; major outer envelope glycoprotein; Gp350; Gp220;
KW   B cell activation; immunoglobulin; secretion; EBV; Gp350/220; antigen;
KW   vaccine; adjuvant; fusion protein; antimicrobial; antiviral; CR2; CD21;
KW   anti-allergic; antitumour; immunostimulatory; infection; tumour;
KW   complement receptor 2.
XX
OS   Human herpesvirus 4.
XX
PN   WO9964603-A2.
XX
PD   16-DEC-1999.
XX
PF   10-JUN-1999;   99WO-US013113.
XX
PR   12-JUN-1998;   98US-0089158P.
XX
PA   (JACK-) JACKSON FOUND ADVANCEMENT MILITARY MED.
PA   (MOND/) MOND J J.
PA   (LEES/) LEES A.
XX
PI   Mond JJ,  Lees A;
XX
DR   WPI; 2000-116546/10.
XX
PT   Adjuvant containing an Epstein-Barr virus sequence that binds to
PT   complement receptor 2, for protecting against, e.g. infection or tumors.
XX
PS   Disclosure; Fig 1A; 43pp; English.
XX
CC   The present invention describes an adjuvant (A) comprising at least one
CC   EBV (Epstein-Barr virus) Gp350/220 sequence (I) that binds to CR2
CC   (complement receptor 2; CD21). Vaccines containing (A) and at least one
CC   other antigenic epitope (also similar immunogens and nucleic acid vectors
CC   that express (I)) are used to elicit an antibody response to both (I)
CC   (for protection against EBV) and the second epitope, which may be
CC   associated with some other infectious agent, allergen, tumour antigen or
CC   generally any condition that requires immune stimulation. The vaccines
CC   and immunogens may also be used to treat B cells in vitro. (A) increases
CC   the immunogenicity of a co-administered antigen. By combining (I), a non-
CC   specific B cell stimulator, with another antigen, the stimulatory effect
CC   is targeted to antigen-specific B cells, i.e. B cells that secrete
CC   immunoglobulins specific for the antigen are stimulated simultaneously
CC   through the antigen receptor and through (I)-mediated crosslinking of
CC   CR2. Constructs of (I) and second antigen are relatively small, so should
CC   be cleared relatively quickly (reducing immunogenicity) and many copies
CC   of second antigen can be included to increase antigenicity. The need for
CC   expensive anti-CR antibodies and potentially inflammatory C3 components
CC   is avoided and folding of the CR2-binding domain in (I)-antigen fusion
CC   proteins is unlikely to be disrupted. The present sequence represents EBV
CC   major outer envelope glycoprotein Gp350. (Updated on 06-AUG-2003 to
CC   correct OS field.)
XX
SQ   Sequence 907 AA;

  Query Match             10.6%;  Score 127.5;  DB 3;  Length 907;
  Best Local Similarity   24.9%;  Pred. No. 0.2;
  Matches   64;  Conservative   26;  Mismatches   88;  Indels   79;  Gaps   14;

Qy          2 TQPASSRVVFDP--SNPT---TFSVEAIAAYTPVALIRLLNASGPLQPGHRVDIADARSI 56
              | | |:     |  ::||   |    |:   ||       ||:||
Db        577 TSPTSAVTTPTPNATSPTLGKTSPTSAVTTPTP-------NATGP-------------- 614

Qy         57 YTVGAAASAARARANHNANTIRRTAMFAETDPMTWLRPTVGLKRTFNPRIIRPQPPNPSM 116
               |||  :  | |  ||   |:  |:            ||          ::  || | :
Db        615 -TVGETSPQANA-TNH---TLGGTS-----------PT---------PVVTSQPKNATS 648

Qy        117 SLGISGPTILPQKTQSADQSALQQPAAL--AFSGSSPQHPP------------PQTTSA 161
              ::      |    | |         |  |  : | :|  | |              | | |
Db        649 AVTTGQHNITSSSTSSMSLRPSSNPETLSPSTSDNSTSHMPLLTSAHPTGGENITQVTPA 708

Qy        162 SVGQQQHVVSGSSGQQPQQGAQSSTVQP---TTGSPPAAQGVPQSTPPPTQNTPQGGKGQ 218
              |:    | || ||  :|: |  |     |  :| : |    | : |||    :||   ||
Db        709 SI--STHHVSTSS-PEPRPGTTSQASGPGNSSTSTKPGEVNVTKGTPPQNATSPQAPSGQ 765

Qy        219 -----TLSHTGQSGNAS 230
                  |:: ||   |::
Db        766 KTAVPTVTSTGGKANST 782


RESULT 7
ABO84900
ID   ABO84900 standard; protein; 634 AA.
XX
AC   ABO84900;
XX
DT   18-NOV-2004  (first entry)
XX
DE   Human cancer-associated protein (CAP) HP07-062.
XX
KW   Human; cancer-associated protein; CAP; cancer; cytostatic.
XX
OS   Homo sapiens.
XX
PN   WO2004058146-A2.
XX
PD   15-JUL-2004.
XX
PF   15-DEC-2003; 2003WO-US040081.
XX
PR   17-DEC-2002; 2002US-00322281.
XX
PA   (SAGR-) SAGRES DISCOVERY INC.
XX
PI   Morris DW,  Malandro MS;
XX
DR   WPI; 2004-499109/47.
DR   N-PSDB; ABD33344.
XX
```

PT Novel human cancer associated protein encoded within open reading frame
PT of cancer associated gene, useful as targets for diagnosing cancer.
XX
PS Claim 18; SEQ ID NO 428; 182pp; English.
XX
CC The invention relates to cancer-associated proteins (CAP) and the cancer-
CC associated (CA) nucleic acids encoding them. The invention also relates
CC to a method for treating cancers involving administering to a patient an
CC inhibitor of CAP, and a method of screening for anticancer activity in a
CC potential drug involving providing a cell that expresses a CA gene,
CC contacting a tissue sample derived from a cancer cell with an anticancer
CC drug candidate and monitoring the effect of the anticancer drug candidate
CC on expression of the CA gene. The CAP proteins are useful for detecting
CC cancer associated with expression of a CAP protein in a test cell sample
CC and for screening for a bioactive agent capable of modulating the
CC activity of a CAP protein. The CA nucleic acids are useful for diagnosing
CC cancer, involving determining the expression of a CA nucleic acid in a
CC tissue. This sequence represents a human CAP of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 634 AA;

Query Match 10.5%; Score 127; DB 8; Length 634;
Best Local Similarity 26.3%; Pred. No. 0.15;
Matches 68; Conservative 17; Mismatches 106; Indels 68; Gaps 12;

```
Qy          1 MTQPASSRVVFDPSNPTTFSVEAIAAYTPVALIRLLNASGPLQP------GHRVDIADAR 54
              |:||        |: | :|||       |     :  :| || |      | :  :| |:
Db        202 MSQPT-------PAQPQSFSVGQPQPPPPPVGGAVAQSSAPLPPFPGAATGPQPMMAAAQ 254

Qy         55 SIYTVGAAASAARARANHNANTIRRTAMFAETDPMTWLRPTVGLKRTFNP-RIIRPQPPN 113
                 ||          :   |: : ||     |     | ||      | :   |||
Db        255 PSQPQGAGPGGQTLPPTN--VTLAQPAMSLPPQP----GPAVGAPAAQQPQQFAYPQPQI 308

Qy        114 PSMSLGISGPTILPQKTQSADQSALQQPAALAFSGSSPQHP--------PPQTTSASVGQ 165
              |   |       ||   | : ||   |       ||  |       |    :  ||
Db        309 PPGHL-------LP--VQPSGQSEYLQQHVAGLQPPSPAQPSSTGAAASPATAATLPVGT 359

Qy        166 QQHVVS------GSSGQQ---------PQQGAQSSTVQPTTGSP------------PAAQ 198
              |:  |      |:| |          | || | :  |||   |             ||
Db        360 GQNASSVGAQLMGASSQPSEAMAPRTGPAQGGQVAPCQPTGVPPATVGGVVQPCLGPAGA 419

Qy        199 GVPQSTPPPTQNTPQGGKG 217
              | ||| |||      || |
Db        420 GQPQSVPPP----QMGGSG 434
```

RESULT 8
ABO84902
ID ABO84902 standard; protein; 724 AA.
XX
AC ABO84902;
XX
DT 18-NOV-2004 (first entry)
XX
DE Human cancer-associated protein (CAP) HP07-062.2.
XX
KW Human; cancer-associated protein; CAP; cancer; cytostatic.
XX
OS Homo sapiens.
XX
PN WO2004058146-A2.
XX
PD 15-JUL-2004.
XX
PF 15-DEC-2003; 2003WO-US040081.
XX
PR 17-DEC-2002; 2002US-00322281.
XX
PA (SAGR-) SAGRES DISCOVERY INC.
XX
PI Morris DW, Malandro MS;
XX
DR WPI; 2004-499109/47.
DR N-PSDB; ABD33346.
XX
PT Novel human cancer associated protein encoded within open reading frame
PT of cancer associated gene, useful as targets for diagnosing cancer.
XX
PS Claim 18; SEQ ID NO 432; 182pp; English.
XX
CC The invention relates to cancer-associated proteins (CAP) and the cancer-
CC associated (CA) nucleic acids encoding them. The invention also relates
CC to a method for treating cancers involving administering to a patient an
CC inhibitor of CAP, and a method of screening for anticancer activity in a
CC potential drug involving providing a cell that expresses a CA gene,
CC contacting a tissue sample derived from a cancer cell with an anticancer
CC drug candidate and monitoring the effect of the anticancer drug candidate
CC on expression of the CA gene. The CAP proteins are useful for detecting
CC cancer associated with expression of a CAP protein in a test cell sample
CC and for screening for a bioactive agent capable of modulating the
CC activity of a CAP protein. The CA nucleic acids are useful for diagnosing
CC cancer, involving determining the expression of a CA nucleic acid in a
CC tissue. This sequence represents a human CAP of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 724 AA;

Query Match 10.5%; Score 127; DB 8; Length 724;
Best Local Similarity 26.3%; Pred. No. 0.17;
Matches 68; Conservative 17; Mismatches 106; Indels 68; Gaps 12;

```
Qy          1 MTQPASSRVVFDPSNPTTFSVEAIAAYTPVALIRLLNASGPLQP------GHRVDIADAR 54
              |:||        |: | :|||       |     :  :| || |      | :  :| |:
Db        202 MSQPT-------PAQPQSFSVGQPQPPPPPVGGAVAQSSAPLPPFPGAATGPQPMMAAAQ 254

Qy         55 SIYTVGAAASAARARANHNANTIRRTAMFAETDPMTWLRPTVGLKRTFNP-RIIRPQPPN 113
                 ||          :   |: : ||     |     | ||      | :   |||
Db        255 PSQPQGAGPGGQTLPPTN--VTLAQPAMSLPPQP----GPAVGAPAAQQPQQFAYPQPQI 308

Qy        114 PSMSLGISGPTILPQKTQSADQSALQQPAALAFSGSSPQHP--------PPQTTSASVGQ 165
              |   |       ||   | : ||   |       ||  |       |    :  ||
Db        309 PPGHL-------LP--VQPSGQSEYLQQHVAGLQPPSPAQPSSTGAAASPATAATLPVGT 359

Qy        166 QQHVVS------GSSGQQ---------PQQGAQSSTVQPTTGSP------------PAAQ 198
              |:  |      |:| |          | || | :  |||   |             ||
Db        360 GQNASSVGAQLMGASSQPSEAMAPRTGPAQGGQVAPCQPTGVPPATVGGVVQPCLGPAGA 419

Qy        199 GVPQSTPPPTQNTPQGGKG 217
              | ||| |||      || |
Db        420 GQPQSVPPP----QMGGSG 434
```

RESULT 9
ADI82549
ID ADI82549 standard; protein; 780 AA.
XX
AC ADI82549;
XX
DT 22-APR-2004 (first entry)
XX
DE Human modifier of p21 (MP21) protein sequence SeqID115.
XX
KW p21 pathway modulating agent; assay system; MP21; cytostatic;
KW MP21 protein activity modulator; cancer; genetically modified animal;
KW human.
XX
OS Homo sapiens.
XX
PN WO2004005486-A2.
XX
PD 15-JAN-2004.

XX
PF 09-JUL-2003; 2003WO-US021510.
XX
PR 10-JUL-2002; 2002US-0394795P.
PR 07-AUG-2002; 2002US-0401739P.
PR 16-SEP-2002; 2002US-0411010P.
PR 30-DEC-2002; 2002US-0437158P.
XX
PA (EXEL-) EXELIXIS INC.
XX
PI Francis-Lang H, Friedman L, Kidd T, Roche S, Joo DM, Lickteig K;
PI Amundsen CD, Hai B, Zhang H, Adamkewicz JI, Hammonds GR;
XX
DR WPI; 2004-091358/09.
DR N-PSDB; ADI82488.
XX
PT Identifying a candidate p21 pathway modulating agent, useful for treating
PT a disease such as cancer, comprises contacting an assay system comprising
PT a MP21 polypeptide or nucleic acid with a test agent.
XX
PS Example 2; SEQ ID NO 115; 392pp; English.
XX
CC This invention relates to a novel candidate p21 pathway modulating agent
CC by contacting an assay system comprising an MP21 (modifier of p21)
CC polypeptide or nucleic acid with a test agent, where in the absence of
CC the test agent the system provides a reference activity and detecting a
CC test agent-biased activity of the assay system. The invention may be
CC useful for the production of compounds with a cytostatic activity through
CC modulation of MP21 protein activity. The MP21 polypeptide or nucleic acid
CC can be used for identifying MP21 modulating agents useful as therapeutic
CC targets for diagnosing cancer or treating disorders associated with
CC defective or impaired p21 and/or MP21 function. MP21 modulating agents
CC are useful in diagnosis, therapy, for example treating cancer, and
CC pharmaceutical development. The genetically modified animals may be used
CC for in vivo assays to test for activity of a candidate p21 modulating
CC agent, or to further assess the role of MP21 in a p21 pathway process.
CC The present sequence is that of a human MP21 protein which is an
CC orthologue of a Drosophila p21 modifier and which was used in the
CC exemplification of the invention.
XX
SQ Sequence 780 AA;

Query Match 10.5%; Score 127; DB 8; Length 780;
Best Local Similarity 26.3%; Pred. No. 0.19;
Matches 68; Conservative 17; Mismatches 106; Indels 68; Gaps 12;

```
Qy          1 MTQPASSRVVFDPSNPTTFSVEAIAAYTPVALIRLLNASGPLQP------GHRVDIADAR 54
              |:||        |: | :|||       |     :  :| || |      | :  :| |:
Db        258 MSQPT-------PAQPQSFSVGQPQPPPPPVGGAVAQSSAPLPPFPGAATGPQPMMAAAQ 310

Qy         55 SIYTVGAAASAARARANHNANTIRRTAMFAETDPMTWLRPTVGLKRTFNP-RIIRPQPPN 113
                ||          :   |: : ||     |     | ||     | :   |||
Db        311 PSQPQGAGPGGQTLPPTN--VTLAQPAMSLPPQP----GPAVGAPAAQQPQQFAYPQPQI 364

Qy        114 PSMSLGISGPTILPQKTQSADQSALQQPAALAFSGSSPQHP--------PPQTTSASVGQ 165
              |   |       ||   | : ||   |         ||  |        |    :  ||
Db        365 PPGHL-------LP--VQPSGQSEYLQQHVAGLQPPSPAQPSSTGAAASPATAATLPVGT 415

Qy        166 QQHVVS------GSSGQQ---------PQQGAQSSTVQPTTGSP------------PAAQ 198
              |:  |       |:| |           | || | :  |||   |             ||
Db        416 GQNASSVGAQLMGASSQPSEAMAPRTGPAQGGQVAPCQPTGVPPATVGGVVQPCLGPAGA 475

Qy        199 GVPQSTPPPTQNTPQGGKG 217
              | ||| |||       || |
Db        476 GQPQSVPPP----QMGGSG 490
```

RESULT 10
AAU86145
ID AAU86145 standard; protein; 839 AA.
XX
AC AAU86145;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human PRO7168 polypeptide.
XX
KW Human; PRO; benign tumour; malignant tumour; lymphoid malignancy;
KW leukaemia; neuronal disorder; stromal disorder; blastocoelic disorder;
KW inflammatory disorder; immune disorder; angiogenic disorder; cytostatic;
KW neuroprotective.
XX
OS Homo sapiens.
XX
PN WO200153486-A1.
XX
PD 26-JUL-2001.
XX
PF 11-FEB-2000; 2000WO-US003565.
XX
PR 08-MAR-1999; 99WO-US005028.
PR 11-MAR-1999; 99US-0123972P.
PR 11-MAY-1999; 99US-0133459P.
PR 02-JUN-1999; 99WO-US012252.
PR 22-JUN-1999; 99US-0140650P.
PR 22-JUN-1999; 99US-0140653P.
PR 20-JUL-1999; 99US-0144758P.
PR 26-JUL-1999; 99US-0145698P.
PR 28-JUL-1999; 99US-0146222P.
PR 17-AUG-1999; 99US-0149395P.
PR 31-AUG-1999; 99US-0151689P.
PR 01-SEP-1999; 99WO-US020111.
PR 15-SEP-1999; 99WO-US021090.
PR 30-NOV-1999; 99WO-US028313.
PR 01-DEC-1999; 99WO-US028301.
PR 01-DEC-1999; 99WO-US028634.
PR 05-JAN-2000; 2000WO-US000219.
XX
PA (GETH) GENENTECH INC.
XX
PI Ashkenazi AJ, Goddard A, Godowski PJ, Gurney AL, Hillan KJ;
PI Marsters SA, Pan J, Pitti RM, Roy MA, Smith V, Stone DM;
PI Watanabe CK, Wood WI;
XX
DR WPI; 2002-205567/26.
DR N-PSDB; ABK40271.
XX
PT Thirty five nucleic acids encoding PRO polypeptides, useful for treating
PT benign or malignant tumors, leukemias and lymphoid malignancies,
PT inflammatory, angiogenic and immunologic disorders.
XX
PS Claim 61; Fig 36; 302pp; English.
XX
CC The present invention relates to the isolation of novel human PRO
CC polypeptides and the polynucleotide sequences encoding them. The PRO
CC polypeptides, agonists, antagonists or anti-PRO antibodies are useful for
CC treating benign or malignant tumours (e.g. renal, kidney, bladder,
CC breast, etc), leukaemias and lymphoid malignancies, other disorders such
CC as neuronal, glial, astrocytal, hypothalamic, glandular, macrophagal,
CC stromal and blastocoelic disorders, inflammatory, immune and angiogenic
CC disorders. The polynucleotide sequences are also useful in gene therapy.
CC AAU86128-AAU86162 represent the human PRO polypeptides of the invention
XX
SQ Sequence 839 AA;

Query Match 10.5%; Score 126.5; DB 5; Length 839;
Best Local Similarity 25.5%; Pred. No. 0.22;
Matches 76; Conservative 24; Mismatches 97; Indels 101; Gaps 16;

```
Qy         17 TTFSVEAIAA-----YTPVALIR-----------LLNASGPLQP----GHRVDIADARS 55
                |  :|||:||       | :| |           : :|: | ||        : :|  |
Db        338 TQVTVEAVAAAGSPPRFPQSLYRGTVARGAGAGVVVKDAAAPSQPLRIQAQDPEFSDLNS 397

Qy         56 IYT--------------VGAAASAARARA------NHN------ANTIRRTAMFAETDP 88
                |              |    : |:| |       ||      | |:    :  :  |
Db        398 AITYRITNHSHFRMEGEVVLTTTTLAQAGAFYAEVEAHNTVTSGTATTVIEIQVSEQEPP 457
```

```
Qy          89 MTWLRPTVGLKRTFNPRIIRPQPPNPSM------SLGISGP-----TILPQKTQS----- 132
               |    | |   :    :  |:|| ||       | | :||     | |   | |
Db         458 STEAGGTTGPWTSTTSEV--PRPPEPSQGPSTTSSGGGTGPHPPSGTTLRPPTSSTPGGP 515

Qy         133 --ADQSALQQPAALAFSGSSPQHPPPQTT-------SASVGQQQHVVSGSSGQQPQQGAQ 183
                 |: |   |||     | : | | | |:          |    |    || : | |: |
Db         516 PGAENSTSHQPATP--GGDTAQTPKPGTSQPMPPGVGTSTSHQPATPSGGTAQTPEPGTS 573

Qy         184 ---------SSTVQPTTGSPPAAQ----GVPQSTPP---------PTQNTPQGGKGQT 219
                      |:: || |     ||    |  |  ||         ||  || ||  ||
Db         574 QPMPPSMGTSTSHQPATPGGGTAQTPEAGTSQPMPPGMGTSTSHQPT--TPGGGTAQT 629
```

RESULT 11
ABG34056
ID ABG34056 standard; protein; 839 AA.
XX
AC ABG34056;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human Pro peptide #27.
XX
KW Human; PRO; secreted protein; transmembrane protein; genetic disorder;
KW tumour; cancer.
XX
OS Homo sapiens.
XX
PN WO200224888-A2.
XX
PD 28-MAR-2002.
XX
PF 29-AUG-2001; 2001WO-US027099.
XX
PR 01-SEP-2000; 2000US-0229896P.
PR 05-SEP-2000; 2000US-0230621P.
PR 22-SEP-2000; 2000US-0235147P.
PR 10-NOV-2000; 2000WO-US030873.
PR 12-JAN-2001; 2001US-0261878P.
PR 16-JAN-2001; 2001US-0261910P.
PR 16-JAN-2001; 2001US-0261939P.
PR 16-JAN-2001; 2001US-0262150P.
PR 25-JAN-2001; 2001US-0264395P.
PR 02-FEB-2001; 2001US-0266421P.
PR 09-FEB-2001; 2001US-0267623P.
PR 28-FEB-2001; 2001WO-US006520.
PR 09-MAR-2001; 2001US-0274399P.
PR 03-APR-2001; 2001US-0280982P.
PR 04-APR-2001; 2001US-0282129P.
PR 04-APR-2001; 2001US-0282199P.
PR 09-MAY-2001; 2001US-0290589P.
PR 25-MAY-2001; 2001WO-US017092.
PR 01-JUN-2001; 2001WO-US017800.
PR 20-JUN-2001; 2001WO-US019692.
PR 29-JUN-2001; 2001WO-US021066.
PR 09-JUL-2001; 2001WO-US021735.
XX
PA (GETH) GENENTECH INC.
XX
PI Baker KP, Eaton DL, Filvaroff E, Goddard A, Grimaldi JC;
PI Gurney AL, Smith V, Stephan J, Watanabe CK, Wood WI, Zhang Z;
PI Fong S;
XX
DR WPI; 2002-362426/39.
DR N-PSDB; ABK69987.
XX
PT New PRO polypeptides and polynucleotides encoding the polypeptides,
PT useful in gene therapy, chromosome identification, tissue typing, or for
PT genetic analysis of individuals with genetic disorders.
XX
PS Claim 11; Fig 54; 218pp; English.
XX
CC This invention relates to the cDNA and protein sequences of novel
CC secreted and transmembrane polypeptides PRO polypeptides. The invention
CC also comprises a method for producing the proteins of the invention by
CC recombinant means and antibodies specific for the protein of the
CC invention. The antibody may be used for detecting the PRO proteins of the
CC invention and may be used to modify their activity. polynucleotides may
CC be used as hybridisation probes for a cDNA library to isolate the full-
CC length PRO cDNA or to isolate other cDNAs, to construct hybridisation
CC probes for mapping the gene which encodes that PRO and for genetic
CC analysis of individuals with genetic disorders, in assays to identify
CC other proteins or molecules involved in binding reaction, to generate
CC transgenic animals or knock-out animals which in turn are useful in the
CC development and screening of therapeutically useful reagents, for
CC chromosome identification, and tissue typing. The PRO polypeptides are
CC useful in gene therapy, and as molecular weight markers for protein
CC electrophoresis purposes. The sequences may also be used to detect
CC overexpression on PRO polypeptides in cancerous tumours and for screening
CC for differentially expressed genes using microarray technology. The
CC present sequence represents a human PRO protein of the invention
XX
SQ Sequence 839 AA;

Query Match 10.5%; Score 126.5; DB 5; Length 839;
Best Local Similarity 25.5%; Pred. No. 0.22;
Matches 76; Conservative 24; Mismatches 97; Indels 101; Gaps 16;

```
Qy          17 TTFSVEAIAA-----YTPVALIR-----------LLNASGPLQP----GHRVDIADARS 55
               |  :|||:||       | :| |            : :|: | ||       : :|  |
Db         338 TQVTVEAVAAAGSPPRFPQSLYRGTVARGAGAGVVVKDAAAPSQPLRIQAQDPEFSDLNS 397

Qy          56 IYT--------------VGAAASAARARA------NHN------ANTIRRTAMFAETDP 88
               |              |    : |:| |       ||      | |:    :  :  |
Db         398 AITYRITNHSHFRMEGEVVLTTTTLAQAGAFYAEVEAHNTVTSGTATTVIEIQVSEQEPP 457

Qy          89 MTWLRPTVGLKRTFNPRIIRPQPPNPSM------SLGISGP-----TILPQKTQS----- 132
               |    | |   :    :  |:|| ||       | | :||     | |   | |
Db         458 STEAGGTTGPWTSTTSEV--PRPPEPSQGPSTTSSGGGTGPHPPSGTTLRPPTSSTPGGP 515

Qy         133 --ADQSALQQPAALAFSGSSPQHPPPQTT-------SASVGQQQHVVSGSSGQQPQQGAQ 183
                 |: |   |||     | : | | | |:          |    |    || : | |: |
Db         516 PGAENSTSHQPATP--GGDTAQTPKPGTSQPMPPGVGTSTSHQPATPSGGTAQTPEPGTS 573

Qy         184 ---------SSTVQPTTGSPPAAQ----GVPQSTPP---------PTQNTPQGGKGQT 219
                      |:: || |     ||    |  |  ||         ||  || ||  ||
Db         574 QPMPPSMGTSTSHQPATPGGGTAQTPEAGTSQPMPPGMGTSTSHQPT--TPGGGTAQT 629
```

RESULT 12
ADA01322
ID ADA01322 standard; protein; 839 AA.
XX
AC ADA01322;
XX
DT 06-NOV-2003 (first entry)
XX
DE Human PRO polypeptide #27.
XX
KW Human; PRO; secreted polypeptide; transmembrane polypeptide;
KW tumour necrosis factor-alpha; TNF-alpha; blood; chondrocyte cell; tumour;
KW adrenal; lung; colon; breast; prostate; rectum; cervix; liver; cancer;
KW microvascular endothelial cell; endothelial cell tube formation;
KW sports-related joint problem; articular cartilage defect; osteoarthritis;
KW rheumatoid arthritis; osteopathic; antirheumatic; antiarthritic.
XX
OS Homo sapiens.
XX
PN US2003068779-A1.
XX
PD 10-APR-2003.
XX
PF 16-SEP-2002; 2002US-00245107.
XX
PR 09-MAY-2001; 2001US-0290589P.

```
PR   29-AUG-2001; 2001WO-US027099.
PR   18-JUL-2002; 2002US-00197942.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Baker KP,  Eaton DL,  Filvaroff E,  Goddard A,  Grimaldi JC;
PI   Gurney AL,  Smith V,  Stephan JP,  Watanabe CK,  Wood WI,  Zhang Z;
PI   Fong S;
XX
DR   WPI; 2003-625484/59.
DR   N-PSDB; ADA01321.
XX
PT   Novel isolated PRO1313, PRO20080 or PRO21383 polypeptide useful for
PT   stimulating proliferation of human microvascular endothelial cells, and
PT   PRO6018 polypeptide useful for stimulating proliferation of chondrocyte
PT   cells.
XX
PS   Claim 11; Fig 54; 307pp; English.
XX
CC   The invention relates to isolated human PRO polypeptides (secreted and
CC   transmembrane polypeptides) and the polynucleotides encoding them. The
CC   invention also relates to an antibody which specifically binds to a PRO
CC   polypeptide, a method for stimulating the release of tumour necrosis
CC   factor-alpha (TNF-alpha) from human blood, a method for stimulating the
CC   proliferation or differentiation of chondrocyte cells and a method for
CC   detecting the presence of a tumour in a mammal (e.g. adrenal, lung,
CC   colon, breast, prostate, rectal, cervical and liver tumours). The
CC   polynucleotides are useful in molecular biology, including uses as
CC   hybridisation probes, in chromosome and gene mapping, in generating
CC   antisense RNA and DNA and in gene therapy. The polynucleotides may also
CC   be used in preparing PRO polypeptides by recombinant techniques and in
CC   generating either transgenic animals or knock-out animals which are
CC   useful in the development and screening of therapeutically useful
CC   reagents. The PRO polypeptides or antibodies are used in preparing a
CC   medicament for treating a condition responsive to the polypeptides or
CC   antibodies, such as tumours, for stimulating and inhibiting proliferation
CC   of human microvascular endothelial cells, for inducing endothelial cell
CC   tube formation and for treating sports-related joint problems, articular
CC   cartilage defects, osteoarthritis and rheumatoid arthritis. This sequence
CC   represents a human PRO polypeptide of the invention.
XX
SQ   Sequence 839 AA;

  Query Match             10.5%;  Score 126.5;  DB 6;  Length 839;
  Best Local Similarity   25.5%;  Pred. No. 0.22;
  Matches   76;  Conservative   24;  Mismatches   97;  Indels  101;  Gaps   16;

Qy         17 TTFSVEAIAA-----YTPVALIR------------LLNASGPLQP----GHRVDIADARS 55
              |  :|||:||       | :| |            : :|: | ||        : :|  |
Db        338 TQVTVEAVAAAGSPPRFPQSLYRGTVARGAGAGVVVKDAAAPSQPLRIQAQDPEFSDLNS 397

Qy         56 IYT---------------VGAAASAARARA------NHN------ANTIRRTAMFAETDP 88
                |               |    : |:| |        ||      | |:    :  :  |
Db        398 AITYRITNHSHFRMEGEVVLTTTTLAQAGAFYAEVEAHNTVTSGTATTVIEIQVSEQEPP 457

Qy         89 MTWLRPTVGLKRTFNPRIIRPQPPNPSM------SLGISGP-----TILPQKTQS----- 132
               |    | |   :    :  |:|| ||       | | :||     | |   | |
Db        458 STEAGGTTGPWTSTTSEV--PRPPEPSQGPSTTSSGGGTGPHPPSGTTLRPPTSSTPGGP 515

Qy        133 --ADQSALQQPAALAFSGSSPQHPPPQTT-------SASVGQQQHVVSGSSGQQPQQGAQ 183
                |: |   |||     | : | | | |:         |   |    || : | |: |
Db        516 PGAENSTSHQPATP--GGDTAQTPKPGTSQPMPPGVGTSTSHQPATPSGGTAQTPEPGTS 573

Qy        184 ---------SSTVQPTTGSPPAAQ----GVPQSTPP---------PTQNTPQGGKGQT 219
                      |:: || |      ||    |  |  ||         ||  || ||  ||
Db        574 QPMPPSMGTSTSHQPATPGGGTAQTPEAGTSQPMPPGMGTSTSHQPT--TPGGGTAQT 629


RESULT 13
ADA43751
ID   ADA43751 standard; protein; 839 AA.
XX
AC   ADA43751;
XX
DT   20-NOV-2003  (first entry)
XX
DE   Human secreted/transmembrane polypeptide PRO7168.
XX
KW   Human; PRO; secreted protein; transmembrane protein;
KW   endothelial cell tube formation; chondrocyte cell differentiation;
KW   microvascular endothelial cell; tumour; lung tumour; colon tumour;
KW   breast tumour; prostate tumour; rectal tumour; kidney tumour;
KW   liver tumour; cytostatic; vaccine.
XX
OS   Homo sapiens.
XX
PN   US2003064474-A1.
XX
PD   03-APR-2003.
XX
PF   16-SEP-2002; 2002US-00245859.
XX
PR   29-AUG-2001; 2001WO-US027099.
PR   18-JUL-2002; 2002US-00197942.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Baker KP,  Eaton DL,  Filvaroff E,  Goddard A,  Grimaldi JC;
PI   Gurney AL,  Smith V,  Stephan JP,  Watanabe CK,  Wood WI,  Zhang Z;
PI   Fong S;
XX
DR   WPI; 2003-605867/57.
DR   N-PSDB; ADA43750.
XX
PT   New isolated nucleic acid encoding a PRO polypeptide, e.g. PRO20080 or
PT   PRO21383, useful in molecular biology, chromosome and gene mapping, in
PT   generating antisense RNA and DNA, and in gene therapy.
XX
PS   Claim 11; Fig 54; 308pp; English.
XX
CC   The invention relates to an isolated secreted/transmembrane (PRO)
CC   polypeptide, having at least 80% sequence identity to a sequence selected
CC   from any one of the 57 amino acid sequences given in specification, or to
CC   a sequence encoded by a nucleic acid molecule selected from any one of
CC   the nucleic acids deposited under any of the ATCC accession numbers given
CC   in specification, or a sequence having at least 80% identity to PRO
CC   lacking its associated signal peptide, an extracellular domain of PRO
CC   with or without its associated signal peptide. Also included are vectors,
CC   transformed host cells, anti-PRO antibodies, the nucleic acids encoding
CC   PRO, PRO fusion proteins, inducing endothelial cell tube formation (by
CC   administering PRO281, PRO1560, PRO189, PRO4499, PRO6308, PRO6000,
CC   PRO10275, PRO21207, PRO20933 or PRO34274 polypeptide or its agonist) and
CC   an oligonucleotide probe derived from any one of the above nucleotide
CC   sequences. PRO6018 polypeptide is useful for stimulating the
CC   proliferation or differentiation of chondrocyte cells. PRO1313, PRO20080
CC   and PRO21383 polypeptides are useful for stimulating the proliferation of
CC   human microvascular endothelial cells. PRO6071, PRO4487 and PRO6006
CC   polypeptides are useful for inhibiting the proliferation of human
CC   microvascular endothelial cells. PRO polypeptides are useful for
CC   detecting the presence of tumour in a mammal, including tumours of lung,
CC   colon, breast, prostate, rectal, kidney and liver. PRO281, PRO1560,
CC   PRO189, PRO4499, PRO6308, PRO6000, PRO10275, PRO21207, PRO20933 and
CC   PRO34274 polypeptides are useful for inducing endothelial cell tube
CC   formation. PRO or the antibody are useful in the preparation of a
CC   medicament for treating a condition responsive to PRO polypeptide. The
CC   oligonucleotide probes are useful for isolating genomic and cDNA
CC   nucleotide sequences, for measuring or detecting the expression of an
CC   associated gene, and as antisense probes. PRO nucleic acid is useful as a
CC   hybridisation probe, in chromosome and gene mapping, in the generation of
CC   antisense RNA and DNA, and for the preparation PRO polypeptides. PRO and
CC   PRO nucleic acid are useful as therapeutic agents, e.g. vaccines. The
CC   present sequence represents a PRO protein.
XX
SQ   Sequence 839 AA;

  Query Match             10.5%;  Score 126.5;  DB 6;  Length 839;
```

```
  Best Local Similarity   25.5%;  Pred. No. 0.22;
  Matches   76;  Conservative   24;  Mismatches   97;  Indels  101;  Gaps   16;

Qy         17 TTFSVEAIAA-----YTPVALIR-----------LLNASGPLQP----GHRVDIADARS 55
              |  :|||:||       | :| |           : :|: | ||       : :|  |
Db        338 TQVTVEAVAAAGSPPRFPQSLYRGTVARGAGAGVVVKDAAAPSQPLRIQAQDPEFSDLNS 397

Qy         56 IYT--------------VGAAASAARARA------NHN------ANTIRRTAMFAETDP 88
                |              |    : |:| |       ||       | |:    :  :  |
Db        398 AITYRITNHSHFRMEGEVVLTTTTLAQAGAFYAEVEAHNTVTSGTATTVIEIQVSEQEPP 457

Qy         89 MTWLRPTVGLKRTFNPRIIRPQPPNPSM------SLGISGP-----TILPQKTQS----- 132
              |    | |   :    :  |:|| ||       | | :||     | |    | |
Db        458 STEAGGTTGPWTSTTSEV--PRPPEPSQGPSTTSSGGGTGPHPPSGTTLRPPTSSTPGGP 515

Qy        133 --ADQSALQQPAALAFSGSSPQHPPPQTT-------SASVGQQQHVVSGSSGQQPQQGAQ 183
              |: |   |||     | : | | | |:         |    |     || : | |: |
Db        516 PGAENSTSHQPATP--GGDTAQTPKPGTSQPMPPGVGTSTSHQPATPSGGTAQTPEPGTS 573

Qy        184 ---------SSTVQPTTGSPPAAQ----GVPQSTPP---------PTQNTPQGGKGQT 219
                      |:: || |      ||    |  |  ||         ||  || ||  ||
Db        574 QPMPPSMGTSTSHQPATPGGGTAQTPEAGTSQPMPPGMGTSTSHQPT--TPGGGTAQT 629


RESULT 14
ADA43519
ID   ADA43519 standard; protein; 839 AA.
XX
AC   ADA43519;
XX
DT   20-NOV-2003  (first entry)
XX
DE   Human secreted/transmembrane polypeptide PRO7168.
XX
KW   Human; PRO; secreted protein; transmembrane protein;
KW   endothelial cell tube formation; chondrocyte cell differentiation;
KW   microvascular endothelial cell; tumour; lung tumour; colon tumour;
KW   breast tumour; prostate tumour; rectal tumour; kidney tumour;
KW   liver tumour; cytostatic; vaccine.
XX
OS   Homo sapiens.
XX
PN   US2003073196-A1.
XX
PD   17-APR-2003.
XX
PF   18-SEP-2002; 2002US-00246210.
XX
PR   04-APR-2001; 2001US-0282199P.
PR   29-AUG-2001; 2001WO-US027099.
PR   18-JUL-2002; 2002US-00197942.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Baker KP,  Eaton DL,  Filvaroff E,  Goddard A,  Grimaldi JC;
PI   Gurney AL,  Smith V,  Stephan JP,  Watanabe CK,  Wood WI,  Zhang Z;
PI   Fong S;
XX
DR   WPI; 2003-743814/70.
DR   N-PSDB; ADA43518.
XX
PT   New secreted and transmembrane PRO polypeptides, e.g. PRO20080 or
PT   PRO21383 useful for stimulating the proliferation or differentiation of
PT   chondrocyte cells and detecting the presence of a tumor in a mammal.
XX
PS   Claim 11; Fig 54; 307pp; English.
XX
CC   The invention relates to an isolated secreted/transmembrane (PRO)
CC   polypeptide, having at least 80% sequence identity to a sequence selected
CC   from any one of the 57 amino acid sequences given in specification, or to
CC   a sequence encoded by a nucleic acid molecule selected from any one of
CC   the nucleic acids deposited under any of the ATCC accession numbers given
CC   in specification, or a sequence having at least 80% identity to PRO
CC   lacking its associated signal peptide, an extracellular domain of PRO
CC   with or without its associated signal peptide. Also included are vectors,
CC   transformed host cells, anti-PRO antibodies, the nucleic acids encoding
CC   PRO, PRO fusion proteins, inducing endothelial cell tube formation (by
CC   administering PRO281, PRO1560, PRO189, PRO4499, PRO6308, PRO6000,
CC   PRO10275, PRO21207, PRO20933 or PRO34274 polypeptide or its agonist) and
CC   an oligonucleotide probe derived from any one of the above nucleotide
CC   sequences. PRO6018 polypeptide is useful for stimulating the
CC   proliferation or differentiation of chondrocyte cells. PRO1313, PRO20080
CC   and PRO21383 polypeptides are useful for stimulating the proliferation of
CC   human microvascular endothelial cells. PRO6071, PRO4487 and PRO6006
CC   polypeptides are useful for inhibiting the proliferation of human
CC   microvascular endothelial cells. PRO polypeptides are useful for
CC   detecting the presence of tumour in a mammal, including tumours of lung,
CC   colon, breast, prostate, rectal, kidney and liver. PRO281, PRO1560,
CC   PRO189, PRO4499, PRO6308, PRO6000, PRO10275, PRO21207, PRO20933 and
CC   PRO34274 polypeptides are useful for inducing endothelial cell tube
CC   formation. PRO or the antibody are useful in the preparation of a
CC   medicament for treating a condition responsive to PRO polypeptide. The
CC   oligonucleotide probes are useful for isolating genomic and cDNA
CC   nucleotide sequences, for measuring or detecting the expression of an
CC   associated gene, and as antisense probes. PRO nucleic acid is useful as a
CC   hybridisation probe, in chromosome and gene mapping, in the generation of
CC   antisense RNA and DNA, and for the preparation PRO polypeptides. PRO and
CC   PRO nucleic acid are useful as therapeutic agents, e.g. vaccines. The
CC   present sequence represents a PRO protein.
XX
SQ   Sequence 839 AA;

  Query Match             10.5%;  Score 126.5;  DB 6;  Length 839;
  Best Local Similarity   25.5%;  Pred. No. 0.22;
  Matches   76;  Conservative   24;  Mismatches   97;  Indels  101;  Gaps   16;

Qy         17 TTFSVEAIAA-----YTPVALIR-----------LLNASGPLQP----GHRVDIADARS 55
              |  :|||:||       | :| |           : :|: | ||       : :|  |
Db        338 TQVTVEAVAAAGSPPRFPQSLYRGTVARGAGAGVVVKDAAAPSQPLRIQAQDPEFSDLNS 397

Qy         56 IYT--------------VGAAASAARARA------NHN------ANTIRRTAMFAETDP 88
                |              |    : |:| |       ||       | |:    :  :  |
Db        398 AITYRITNHSHFRMEGEVVLTTTTLAQAGAFYAEVEAHNTVTSGTATTVIEIQVSEQEPP 457

Qy         89 MTWLRPTVGLKRTFNPRIIRPQPPNPSM------SLGISGP-----TILPQKTQS----- 132
              |    | |   :    :  |:|| ||       | | :||     | |    | |
Db        458 STEAGGTTGPWTSTTSEV--PRPPEPSQGPSTTSSGGGTGPHPPSGTTLRPPTSSTPGGP 515

Qy        133 --ADQSALQQPAALAFSGSSPQHPPPQTT-------SASVGQQQHVVSGSSGQQPQQGAQ 183
              |: |   |||     | : | | | |:         |    |     || : | |: |
Db        516 PGAENSTSHQPATP--GGDTAQTPKPGTSQPMPPGVGTSTSHQPATPSGGTAQTPEPGTS 573

Qy        184 ---------SSTVQPTTGSPPAAQ----GVPQSTPP---------PTQNTPQGGKGQT 219
                      |:: || |      ||    |  |  ||         ||  || ||  ||
Db        574 QPMPPSMGTSTSHQPATPGGGTAQTPEAGTSQPMPPGMGTSTSHQPT--TPGGGTAQT 629


RESULT 15
ADA01194
ID   ADA01194 standard; protein; 839 AA.
XX
AC   ADA01194;
XX
DT   06-NOV-2003  (first entry)
XX
DE   Human PRO polypeptide #27.
XX
KW   Human; PRO; secreted polypeptide; transmembrane polypeptide;
KW   tumour necrosis factor-alpha; TNF-alpha; blood; chondrocyte cell; tumour;
KW   cancer; lung; colon; breast; prostate; rectum; kidney; liver;
KW   microvascular endothelial cell; endothelial cell tube formation.
XX
OS   Homo sapiens.
XX
PN   US2003068782-A1.
XX
```

PD 10-APR-2003.
XX
PF 16-SEP-2002; 2002US-00245851.
XX
PR 27-APR-1999; 99US-0131271P.
PR 29-OCT-1999; 99US-0162506P.
PR 02-DEC-1999; 99WO-US028551.
PR 29-AUG-2001; 2001WO-US027099.
PR 18-JUL-2002; 2002US-00197942.
XX
PA (GETH) GENENTECH INC.
XX
PI Baker KP, Eaton DL, Filvaroff E, Goddard A, Grimaldi JC;
PI Gurney AL, Smith V, Stephan JP, Watanabe CK, Wood WI, Zhang Z;
PI Fong S;
XX
DR WPI; 2003-625487/59.
DR N-PSDB; ADA01193.
XX
PT Novel isolated PRO polypeptides e.g. PRO281 and PRO1560, useful in the
PT preparation of a medicament for treating a condition responsive to PRO
PT polypeptide, and as therapeutic agents e.g. vaccines.
XX
PS Claim 11; Fig 54; 308pp; English.
XX
CC The invention relates to isolated human PRO polypeptides (secreted and
CC transmembrane polypeptides) and the polynucleotides encoding them. The
CC invention also relates to an antibody which specifically binds to a PRO
CC polypeptide, a method for stimulating the release of tumour necrosis
CC factor-alpha (TNF-alpha) from human blood, a method for stimulating the
CC proliferation or differentiation of chondrocyte cells and a method for
CC detecting the presence of a tumour in a mammal (e.g. lung, colon, breast,
CC prostate, rectal, kidney and liver tumours). The polynucleotides are
CC useful in molecular biology, including uses as hybridisation probes, in
CC chromosome and gene mapping, in generating antisense RNA and DNA and in
CC gene therapy. The polynucleotides may also be used in preparing PRO
CC polypeptides by recombinant techniques and in generating either
CC transgenic animals or knock-out animals which are useful in the
CC development and screening of therapeutically useful reagents. The PRO
CC polypeptides or antibodies are used in preparing a medicament for
CC treating a condition responsive to the polypeptides or antibodies, such
CC as tumours, for stimulating and inhibiting proliferation of human
CC microvascular endothelial cells and for inducing endothelial cell tube
CC formation. This sequence represents a human PRO polypeptide of the
CC invention.
XX
SQ Sequence 839 AA;

Query Match 10.5%; Score 126.5; DB 6; Length 839;
Best Local Similarity 25.5%; Pred. No. 0.22;
Matches 76; Conservative 24; Mismatches 97; Indels 101; Gaps 16;

```
Qy          17 TTFSVEAIAA-----YTPVALIR------------LLNASGPLQP----GHRVDIADARS 55
               |  :|||:||       | :| |             : :|: | ||        : :|  |
Db         338 TQVTVEAVAAAGSPPRFPQSLYRGTVARGAGAGVVVKDAAAPSQPLRIQAQDPEFSDLNS 397

Qy          56 IYT---------------VGAAASAARARA------NHN------ANTIRRTAMFAETDP 88
                |               |    : |:| |       ||      | |:    :  :  |
Db         398 AITYRITNHSHFRMEGEVVLTTTTLAQAGAFYAEVEAHNTVTSGTATTVIEIQVSEQEPP 457

Qy          89 MTWLRPTVGLKRTFNPRIIRPQPPNPSM------SLGISGP-----TILPQKTQS----- 132
               |    | |   :    :  |:|| ||       | | :||     | |   | |
Db         458 STEAGGTTGPWTSTTSEV--PRPPEPSQGPSTTSSGGGTGPHPPSGTTLRPPTSSTPGGP 515

Qy         133 --ADQSALQQPAALAFSGSSPQHPPPQTT-------SASVGQQQHVVSGSSGQQPQQGAQ 183
                 |: |   |||      | : | | | |:         |   |     || : | |: |
Db         516 PGAENSTSHQPATP--GGDTAQTPKPGTSQPMPPGVGTSTSHQPATPSGGTAQTPEPGTS 573

Qy         184 ---------SSTVQPTTGSPPAAQ----GVPQSTPP---------PTQNTPQGGKGQT 219
                      |:: || |      ||     |  |  ||         ||  || ||  ||
Db         574 QPMPPSMGTSTSHQPATPGGGTAQTPEAGTSQPMPPGMGTSTSHQPT--TPGGGTAQT 629
```

Search completed: September 25, 2005, 20:55:23
Job time : 328 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 25, 2005, 18:22:58 ; Search time 389 Seconds
(without alignments)
309.354 Million cell updates/sec

Title: US-09-874-140-2
Perfect score: 1206
Sequence: 1 MTQPASSRVVFDPSNPTTFS.....KGQTLSHTGQSGNASRSRRV 235

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1206	100.0	235	1	VP26_VZVD	P09279 varicella-z
2	1206	100.0	235	2	Q6QCN2	Q6qcn2 human herpe
3	473.5	39.3	228	2	Q9E1Z3	Q9e1z3 cercopithec
4	175	14.5	124	2	Q65554	Q65554 bovine herp
5	175	14.5	124	2	Q77CD3	Q77cd3 bovine herp
6	174	14.4	103	2	Q77KA6	Q77ka6 suid herpes
7	174	14.4	103	2	Q9ICS6	Q9ics6 pseudorabie
8	173	14.3	125	2	Q6X247	Q6x247 bovine herp
9	154.5	12.8	119	2	O39267	O39267 equid herpe
10	153.5	12.7	119	1	VP26_EHV1B	P28974 equine herp
11	153.5	12.7	119	2	Q6S6P6	Q6s6p6 equid herpe
12	142	11.8	503	2	Q66IV2	Q66iv2 xenopus lae
13	141.5	11.7	771	2	Q76N76	Q76n76 rattus norv
14	139	11.5	1114	2	Q6P9L3	Q6p9l3 mus musculu
15	138	11.4	776	2	Q9JJ50	Q9jj50 rattus norv
16	133.5	11.1	508	1	ABI1_HUMAN	Q8izp0 homo sapien
17	132.5	11.0	165	2	Q9DPQ4	Q9dpq4 meleagrid h
18	132.5	11.0	237	2	Q9E1F9	Q9e1f9 meleagrid h
19	132	10.9	862	1	MCDL_RAT	Q9jik1 rattus norv
20	131.5	10.9	2232	2	Q8IFX6	Q8ifx6 caenorhabdi
21	131	10.9	530	1	NLFA_MOUSE	Q8bg30 mus musculu
22	129.5	10.7	686	2	Q7S0F3	Q7s0f3 neurospora
23	128	10.6	1151	2	O57580	O57580 gallus gall
24	127.5	10.6	907	2	Q66537	Q66537 human herpe
25	127	10.5	721	1	YK82_MYCTU	Q10690 mycobacteri
26	127	10.5	721	2	Q7TZ24	Q7tz24 mycobacteri
27	127	10.5	780	2	Q6PI50	Q6pi50 homo sapien
28	127	10.5	784	2	Q90YB5	Q90yb5 gallus gall
29	127	10.5	825	2	O75157	O75157 homo sapien
30	127	10.5	829	2	Q6N5Q1	Q6n5q1 rhodopseudo
31	126.5	10.5	2321	2	Q6R5R1	Q6r5r1 meleagrid h
32	126.5	10.5	2321	2	Q9DGT6	Q9dgt6 meleagrid h
33	126.5	10.5	2323	2	Q6HAA3	Q6haa3 meleagrid h
34	126	10.4	775	2	Q99LI8	Q99li8 mus musculu
35	126	10.4	775	2	Q61691	Q61691 mus musculu
36	125.5	10.4	70	2	Q8UZ10	Q8uz10 pseudorabie
37	125.5	10.4	647	2	Q9VQF3	Q9vqf3 drosophila
38	125.5	10.4	760	1	HRS_DROME	Q960x8 drosophila
39	125.5	10.4	783	2	O91331	O91331 cercopithec
40	125.5	10.4	800	2	Q6CGD7	Q6cgd7 yarrowia li
41	125.5	10.4	2169	2	Q8NPN6	Q8npn6 corynebacte
42	125	10.4	1061	2	Q9V853	Q9v853 drosophila
43	125	10.4	1724	2	P91019	P91019 caenorhabdi
44	124.5	10.3	131	2	Q9E6N4	Q9e6n4 meleagrid h
45	124.5	10.3	513	2	Q9FTT8	Q9ftt8 oryza sativ

ALIGNMENTS

RESULT 1
VP26_VZVD
ID VP26_VZVD STANDARD; PRT; 235 AA.
AC P09279;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Capsid protein VP26.
GN Name=23;
OS Varicella-zoster virus (strain Dumas) (VZV).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=10338;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86306657; PubMed=3018124;
RA Davison A.J., Scott J.E.;
RT "The complete DNA sequence of varicella-zoster virus.";
RL J. Gen. Virol. 67:1759-1816(1986).
CC -!- SIMILARITY: Belongs to the herpesviruses UL35 family.
CC ----------
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC ----------
DR EMBL; X04370; CAA27906.1; -.
DR PIR; E27343; WZBE23.
DR InterPro; IPR007584; UL35.
DR Pfam; PF04496; Herpes_UL35; 1.
KW Coat protein.
SQ SEQUENCE 235 AA; 24417 MW; 6BC37A7BEE06F30A CRC64;

Query Match 100.0%; Score 1206; DB 1; Length 235;
Best Local Similarity 100.0%; Pred. No. 9.5e-64;
Matches 235; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MTQPASSRVVFDPSNPTTFSVEAIAAYTPVALIRLLNASGPLQPGHRVDIADARSIYTVG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MTQPASSRVVFDPSNPTTFSVEAIAAYTPVALIRLLNASGPLQPGHRVDIADARSIYTVG 60

Qy         61 AAASAARARANHNANTIRRTAMFAETDPMTWLRPTVGLKRTFNPRIIRPQPPNPSMSLGI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AAASAARARANHNANTIRRTAMFAETDPMTWLRPTVGLKRTFNPRIIRPQPPNPSMSLGI 120

Qy        121 SGPTILPQKTQSADQSALQQPAALAFSGSSPQHPPPQTTSASVGQQQHVVSGSSGQQPQQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SGPTILPQKTQSADQSALQQPAALAFSGSSPQHPPPQTTSASVGQQQHVVSGSSGQQPQQ 180

Qy        181 GAQSSTVQPTTGSPPAAQGVPQSTPPPTQNTPQGGKGQTLSHTGQSGNASRSRRV 235
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        181 GAQSSTVQPTTGSPPAAQGVPQSTPPPTQNTPQGGKGQTLSHTGQSGNASRSRRV 235


RESULT 2
Q6QCN2
ID   Q6QCN2      PRELIMINARY;      PRT;   235 AA.
AC   Q6QCN2;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Capsid protein.
OS   Human herpesvirus 3 (HHV-3) (Varicella-zoster virus).
OC   Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC   Alphaherpesvirinae; Varicellovirus.
OX   NCBI_TaxID=10335;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BC, and MSP;
RA   Grose C., Tyler S., Peters G., Hiebert J., Stephens G., Ruyechan W.,
RA   Jackson W., Storlie J., Tipples G.A.;
RT   "Complete DNA sequence analyses of the first two gE (D150N) VZV mutant
RT   viruses in North America: Genotypes with an accelerated cell spread
RT   phenotype.";
RL   J. Virol. 0:0-0(2004).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MSP;
RX   MEDLINE=20455576; PubMed=10998331; DOI=10.1006/viro.2000.0507;
RA   Santos R.A., Hatfield C.C., Cole N.L., Padilla J.A., Moffat J.F.,
RA   Arvin A.M., Ruyechan W.T., Hay J., Grose C.;
RT   "Varicella-zoster virus gE escape mutant VZV-MSP exhibits an
RT   accelerated cell-to-cell spread phenotype in both infected cell
RT   cultures and SCID-hu mice.";
RL   Virology 275:306-317(2000).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MSP;
RA   Santos R.A., Hatfield C.C., Cole N.L., Padilla J.A., Moffat J.F.,
RA   Arvin A.M., Ruyechan W.T., Hay J., Grose C.;
RL   Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MSP;
RA   Grose C., Tyler S., Peters G., Hiebert J., Stephens G., Ruyechan W.,
RA   Jackson W., Storlie J., Tipples G.A.;
RL   Submitted (FEB-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY548171; AAT07705.1; -.
DR   EMBL; AY548170; AAT07781.1; -.
DR   GO; GO:0019028; C:viral capsid; IEA.
DR   InterPro; IPR007584; UL35.
DR   Pfam; PF04496; Herpes_UL35; 1.
SQ   SEQUENCE   235 AA;  24417 MW;  6BC37A7BEE06F30A CRC64;

  Query Match             100.0%;  Score 1206;  DB 2;  Length 235;
  Best Local Similarity   100.0%;  Pred. No. 9.5e-64;
  Matches  235;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MTQPASSRVVFDPSNPTTFSVEAIAAYTPVALIRLLNASGPLQPGHRVDIADARSIYTVG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MTQPASSRVVFDPSNPTTFSVEAIAAYTPVALIRLLNASGPLQPGHRVDIADARSIYTVG 60

Qy         61 AAASAARARANHNANTIRRTAMFAETDPMTWLRPTVGLKRTFNPRIIRPQPPNPSMSLGI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AAASAARARANHNANTIRRTAMFAETDPMTWLRPTVGLKRTFNPRIIRPQPPNPSMSLGI 120

Qy        121 SGPTILPQKTQSADQSALQQPAALAFSGSSPQHPPPQTTSASVGQQQHVVSGSSGQQPQQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SGPTILPQKTQSADQSALQQPAALAFSGSSPQHPPPQTTSASVGQQQHVVSGSSGQQPQQ 180

Qy        181 GAQSSTVQPTTGSPPAAQGVPQSTPPPTQNTPQGGKGQTLSHTGQSGNASRSRRV 235
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GAQSSTVQPTTGSPPAAQGVPQSTPPPTQNTPQGGKGQTLSHTGQSGNASRSRRV 235


RESULT 3
Q9E1Z3
ID   Q9E1Z3      PRELIMINARY;      PRT;   228 AA.
AC   Q9E1Z3;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Hypothetical protein.
OS   Cercopithecine herpesvirus 7.
OC   Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC   Alphaherpesvirinae; Varicellovirus.
OX   NCBI_TaxID=35245;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21251382; PubMed=11352673; DOI=10.1006/viro.2001.0912;
RA   Gray W.L., Starnes B., White M.W., Mahalingam R.;
RT   "The DNA sequence of the simian varicella virus genome.";
RL   Virology 284:123-130(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Gray W.L., Starnes H.B., White M.W., Ashburn C.V., Mahalingam R.;
RL   Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF275348; AAG27196.1; -.
DR   GO; GO:0019028; C:viral capsid; IEA.
DR   InterPro; IPR007584; UL35.
DR   Pfam; PF04496; Herpes_UL35; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   228 AA;  23611 MW;  3A1F4512E4426E37 CRC64;

  Query Match             39.3%;  Score 473.5;  DB 2;  Length 228;
  Best Local Similarity   47.5%;  Pred. No. 1.1e-20;
  Matches  122;  Conservative   21;  Mismatches   63;  Indels   51;  Gaps    8;

Qy          1 MTQPASSRVVFDPSNPTTFSVEAIAAYTPVALIRLLNASGPLQPGHRVDIADARSIYTVG 60
              |||  |||  |||| | ||||:|||| :||||::|:|||||:  | | ||   |||||||
Db          1 MTQSTSSRAHFDPSVPGTFSVDAIAANSPVALLQLVNASGPVISGVRFDIGHIRSIYTVG 60

Qy         61 AAASAARARANHNANTIRRTAMFAETDPMTWLRPTVGLKRTFNPRIIRPQPPNPSMSLGI 120
              |||:|||:| ||| |||||||||||| ||||||||||||||||:|||     :|:
Db         61 AAAAAARSRVNHNVNTIRRTAMFAETDAMTWLRPTVGLKRTFNPRIVRPQ-----VSVIN 115

Qy        121 SGPTILPQKTQSADQSALQQPAAL-----AFSGSSPQHP-------PPQTTSASVGQQQH 168
              |   |: : | |  |        |:         |: ||       | |:|  :   |
Db        116 SQANIIQESTNSVSQQPPPSNQAVHSHVSILPSSTSQHSNQVNANIPAQSTQVTAPGQSD 175

Qy        169 VVSGSSGQQPQQGAQS----------STVQPTTGSPPAAQGVPQSTPPPTQNTPQGGKGQ 218
                ||    ||   ||:          :||  |:||   ||  | |          ||:|
Db        176 QQSG----QPAVSAQATVANTGVSAQATVANTSGSQSKAQ--PAS----------GGQG- 218

Qy        219 TLSHTGQSGNASRSRRV 235
                  ||    ||||
Db        219 -------SGTPRSSRRV 228


RESULT 4
Q65554
ID   Q65554      PRELIMINARY;      PRT;   124 AA.
AC   Q65554;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   UL35.
OS   Bovine herpesvirus 1.
OC   Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC   Alphaherpesvirinae; Varicellovirus.
OX   NCBI_TaxID=10320;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Cooper;
RA   Schwyzer M., Vlcek C., Lowery D.E., Bello L.J., Meyer G., Misra V.,
RA   Thiry E., Paces V.;
RL   Submitted (AUG-1996) to the EMBL/GenBank/DDBJ databases.
```

```
DR   EMBL; Z78205; CAB01606.1; -.
DR   GO; GO:0019028; C:viral capsid; IEA.
DR   InterPro; IPR007584; UL35.
DR   Pfam; PF04496; Herpes_UL35; 1.
SQ   SEQUENCE   124 AA;  13420 MW;  31FF6C0DFA43CD81 CRC64;

  Query Match             14.5%;  Score 175;  DB 2;  Length 124;
  Best Local Similarity   36.3%;  Pred. No. 0.0022;
  Matches   41;  Conservative   18;  Mismatches   48;  Indels    6;  Gaps    3;

Qy          1 MTQPASSRVVFDPSNPTTFSVEAIAAYTPVALIRLLNASG-PLQPGHRVD-IADARSIYT 58
              |: ||      ||| | | : : :    || :: :|||:  ||| |   : :: ||
Db          1 MSAPAG----IDPSAPATITPDTLRDLLPVQILHVLNAAARPLQDGTTPEQVSAARRNLL 56

Qy         59 VGAAASAARARANHNANTIRRTAMFAETDPMTWLRPTVGLKRTFNPRIIRPQP 111
              || : :    | |    : |  ||| |   | |||:|||||| ||::  |
Db         57 VGTSLAMVDLRRRHEKAVVPRVPMFATYDHAHWARPTIGLKRTFLPRVVQLPP 109


RESULT 5
Q77CD3
ID   Q77CD3      PRELIMINARY;      PRT;   124 AA.
AC   Q77CD3;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Capsid protein.
GN   Name=UL35;
OS   Bovine herpesvirus type 1.1.
OC   Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC   Alphaherpesvirinae; Varicellovirus.
OX   NCBI_TaxID=79889;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Schwyzer M., Vlcek C., Lowery D.E., Bello L.J., Meyer G., Misra V.;
RL   Submitted (SEP-1997) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Schwyzer M., Paces V., Letchworth G.J., Misra V., Buhk H.J.,
RA   Lowery D.E., Simard C., Bello L.J., Thiry E., Vlcek C.;
RL   Submitted (SEP-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ004801; CAA06098.1; -.
DR   GO; GO:0019028; C:viral capsid; IEA.
DR   InterPro; IPR007584; UL35.
DR   Pfam; PF04496; Herpes_UL35; 1.
SQ   SEQUENCE   124 AA;  13420 MW;  31FF6C0DFA43CD81 CRC64;

  Query Match             14.5%;  Score 175;  DB 2;  Length 124;
  Best Local Similarity   36.3%;  Pred. No. 0.0022;
  Matches   41;  Conservative   18;  Mismatches   48;  Indels    6;  Gaps    3;

Qy          1 MTQPASSRVVFDPSNPTTFSVEAIAAYTPVALIRLLNASG-PLQPGHRVD-IADARSIYT 58
              |: ||      ||| | | : : :    || :: :|||:  ||| |   : :: ||
Db          1 MSAPAG----IDPSAPATITPDTLRDLLPVQILHVLNAAARPLQDGTTPEQVSAARRNLL 56

Qy         59 VGAAASAARARANHNANTIRRTAMFAETDPMTWLRPTVGLKRTFNPRIIRPQP 111
              || : :    | |    : |  ||| |   | |||:|||||| ||::  |
Db         57 VGTSLAMVDLRRRHEKAVVPRVPMFATYDHAHWARPTIGLKRTFLPRVVQLPP 109


RESULT 6
Q77KA6
ID   Q77KA6      PRELIMINARY;      PRT;   103 AA.
AC   Q77KA6;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   UL35 protein.
GN   Name=UL35;
OS   Suid herpesvirus 1.
OC   Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC   Alphaherpesvirinae; Varicellovirus.
OX   NCBI_TaxID=10345;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Becker;
RX   MEDLINE=21145867; PubMed=11248101; DOI=10.1073/pnas.061029798;
RA   Smith G.A., Gross S.P., Enquist L.W.;
RT   "Herpesviruses use bidirectional fast-axonal transport to spread in
RT   sensory neurons.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:3466-3470(2001).
DR   EMBL; AF301599; AAG21381.1; -.
DR   GO; GO:0019028; C:viral capsid; IEA.
DR   InterPro; IPR007584; UL35.
DR   Pfam; PF04496; Herpes_UL35; 1.
SQ   SEQUENCE   103 AA;  11474 MW;  0B8757DB8B7B0ECC CRC64;

  Query Match             14.4%;  Score 174;  DB 2;  Length 103;
  Best Local Similarity   37.1%;  Pred. No. 0.0021;
  Matches   39;  Conservative   17;  Mismatches   41;  Indels    8;  Gaps    3;

Qy         11 FDPSNPTTFSVEAIAAYTPVALIRLLNASGPLQ----PGHRVDIADARSIYTVGAAASAA 66
              |||:|| | : : :    || ::  || :  ||      | : : |||    :| : :
Db          3 FDPNNPRTITAQTLEGALPVDILLRLNRATGLQMDAAEAHAI-VEDARRTLFIGTSLALV 61

Qy         67 RARANHNANTIRRTAMFAETDPMTWLRPTVGLKRTFNPRIIRPQP 111
              |  |: : : |  ||| :|  :| |||||||||| |   || |
Db         62 NLRHAHDKHLVERQPMFATSDYSSWARPTVGLKRTFCP---RPPP 103


RESULT 7
Q9ICS6
ID   Q9ICS6      PRELIMINARY;      PRT;   103 AA.
AC   Q9ICS6;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   UL35 protein.
GN   Name=UL35;
OS   Pseudorabies virus (strain Kaplan) (PRV).
OC   Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC   Alphaherpesvirinae; Varicellovirus.
OX   NCBI_TaxID=33703;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Kaplan;
RX   MEDLINE=20481646; PubMed=11024135;
RX   DOI=10.1128/JVI.74.21.10063-10073.2000;
RA   Klupp B.G., Granzow H., Mettenleiter T.C.;
RT   "Primary envelopment of pseudorabies virus at the nuclear membrane
RT   requires the UL34 gene product.";
RL   J. Virol. 74:10063-10073(2000).
DR   EMBL; AJ276165; CAB98182.1; -.
DR   GO; GO:0019028; C:viral capsid; IEA.
DR   InterPro; IPR007584; UL35.
DR   Pfam; PF04496; Herpes_UL35; 1.
SQ   SEQUENCE   103 AA;  11474 MW;  0B8757DB8B7B0ECC CRC64;

  Query Match             14.4%;  Score 174;  DB 2;  Length 103;
  Best Local Similarity   37.1%;  Pred. No. 0.0021;
  Matches   39;  Conservative   17;  Mismatches   41;  Indels    8;  Gaps    3;

Qy         11 FDPSNPTTFSVEAIAAYTPVALIRLLNASGPLQ----PGHRVDIADARSIYTVGAAASAA 66
              |||:|| | : : :    || ::  || :  ||      | : : |||    :| : :
Db          3 FDPNNPRTITAQTLEGALPVDILLRLNRATGLQMDAAEAHAI-VEDARRTLFIGTSLALV 61

Qy         67 RARANHNANTIRRTAMFAETDPMTWLRPTVGLKRTFNPRIIRPQP 111
              |  |: : : |  ||| :|  :| |||||||||| |   || |
Db         62 NLRHAHDKHLVERQPMFATSDYSSWARPTVGLKRTFCP---RPPP 103


RESULT 8
Q6X247
ID   Q6X247      PRELIMINARY;      PRT;   125 AA.
AC   Q6X247;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
```

```
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   UL35 capsid protein.
GN   ORFNames=BHV5-21;
OS   Bovine herpesvirus 5.
OC   Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC   Alphaherpesvirinae; Varicellovirus.
OX   NCBI_TaxID=35244;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SV507/99;
RX   MEDLINE=22850801; PubMed=12970418;
RX   DOI=10.1128/JVI.77.19.10339-10347.2003;
RA   Delhon G., Moraes M.P., Lu Z., Afonso C.L., Flores E.F., Weiblen R.,
RA   Kutish G.F., Rock D.L.;
RT   "Genome of bovine herpesvirus 5.";
RL   J. Virol. 77:10339-10347(2003).
DR   EMBL; AY261359; AAR86126.1; -.
DR   GO; GO:0019028; C:viral capsid; IEA.
DR   InterPro; IPR007584; UL35.
DR   Pfam; PF04496; Herpes_UL35; 1.
SQ   SEQUENCE   125 AA;  13525 MW;  BC5634AF35F5289F CRC64;

  Query Match             14.3%;  Score 173;  DB 2;  Length 125;
  Best Local Similarity   37.3%;  Pred. No. 0.0029;
  Matches   41;  Conservative   17;  Mismatches   46;  Indels    6;  Gaps    3;

Qy          1 MTQPASSRVVFDPSNPTTFSVEAIAAYTPVALIRLLN-ASGPLQPGHRVD-IADARSIYT 58
              |: ||      ||| | | : : :    || :: :|| || ||| |   : :: ||
Db          1 MSAPAG----IDPSAPATITPDTLRDLLPVQILHVLNAASRPLQDGTTPEQVSAARRNLL 56

Qy         59 VGAAASAARARANHNANTIRRTAMFAETDPMTWLRPTVGLKRTFNPRIIR 108
              || : :    |  |    : |  |||  |   | |||:|||||| ||::
Db         57 VGTSLAMVDLRRRHEKAVVPRVPMFATYDHAHWARPTIGLKRTFLPRVVQ 106


RESULT 9
O39267
ID   O39267      PRELIMINARY;      PRT;   119 AA.
AC   O39267;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   25.
OS   Equid herpesvirus 4.
OC   Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC   Alphaherpesvirinae; Varicellovirus.
OX   NCBI_TaxID=10331;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NS80567;
RX   MEDLINE=98264497; PubMed=9603335;
RA   Telford E.A., Watson M.S., Perry J., Cullinane A.A., Davison A.J.;
RT   "The DNA sequence of equine herpesvirus-4.";
RL   J. Gen. Virol. 79:1197-1203(1998).
DR   EMBL; AF030027; AAC59540.1; -.
DR   PIR; T42568; T42568.
DR   GO; GO:0019028; C:viral capsid; IEA.
DR   InterPro; IPR007584; UL35.
DR   Pfam; PF04496; Herpes_UL35; 1.
SQ   SEQUENCE   119 AA;  13667 MW;  CC6E4FDD70E2E55E CRC64;

  Query Match             12.8%;  Score 154.5;  DB 2;  Length 119;
  Best Local Similarity   32.8%;  Pred. No. 0.034;
  Matches   38;  Conservative   12;  Mismatches   39;  Indels   27;  Gaps    3;

Qy          9 VVFDPSNPTTFSVEAIAAYTPVALIRLLNASGPLQPGHRVDIADARSIYT---------- 58
              | ||| :||           || :: :||          :|  |    ||
Db         12 VAFDPWDPTNIKAANFKDMLPVDVMTILN--------QNIDELD----YTKYSDDEINEG 59

Qy         59 -----VGAAASAARARANHNANTIRRTAMFAETDPMTWLRPTVGLKRTFNPRIIRP 109
                  :| | :    |  |    :||:||||  |  || || :|||||| || ::|
Db         60 LKQLFMGTAKTMVVLRQRHLKALVRRSAMFAHNDASTWARPNIGLKRTFPPRFMQP 115


RESULT 10
VP26_EHV1B
ID   VP26_EHV1B     STANDARD;      PRT;   119 AA.
AC   P28974;
DT   01-DEC-1992 (Rel. 24, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Capsid protein VP26.
GN   Name=25;
OS   Equine herpesvirus 1 (strain Ab4p) (EHV-1) (Equine abortion virus).
OC   Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC   Alphaherpesvirinae; Varicellovirus.
OX   NCBI_TaxID=31520;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92295566; PubMed=1318606;
RA   Telford E.A.R., Watson M.S., McBride K., Davison A.J.;
RT   "The DNA sequence of equine herpesvirus-1.";
RL   Virology 189:304-316(1992).
CC   -!- SIMILARITY: Belongs to the herpesviruses UL35 family.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; M86664; AAB02460.1; -.
DR   PIR; H36797; WZBEB7.
DR   InterPro; IPR007584; UL35.
DR   Pfam; PF04496; Herpes_UL35; 1.
KW   Coat protein.
SQ   SEQUENCE   119 AA;  13596 MW;  3603A1AA4208CB08 CRC64;

  Query Match             12.7%;  Score 153.5;  DB 1;  Length 119;
  Best Local Similarity   31.1%;  Pred. No. 0.039;
  Matches   37;  Conservative   16;  Mismatches   45;  Indels   21;  Gaps    3;

Qy          3 QPASSRVVFDPSNPTTFSVEAIAAYTPVALIRLLNASGPLQPGHRVDIADARSIYT---- 58
              |   : | |:|::|            || :| :||          :|  |    ||
Db          6 QQQQAPVAFNPADPPNIKAANFKDMLPVDVITILN--------QNIDELDYTK-YTEDEI 56

Qy         59 --------VGAAASAARARANHNANTIRRTAMFAETDPMTWLRPTVGLKRTFNPRIIRP 109
                     :| | :    |  |  : :||: |||: |  || || :|||||| || ::|
Db         57 SEGLKQLFMGTARTMVSLRQRHLKSLVRRSDMFAQNDASTWARPNIGLKRTFPPRFMQP 115


RESULT 11
Q6S6P6
ID   Q6S6P6      PRELIMINARY;      PRT;   119 AA.
AC   Q6S6P6;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Capsid protein.
OS   Equid herpesvirus 1.
OC   Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC   Alphaherpesvirinae; Varicellovirus.
OX   NCBI_TaxID=10326;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=V592;
RA   Davis-Poynter N.J., Nugent J., Birch-Machin I., Allen G.P.;
RL   Submitted (NOV-2003) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Ab4;
RX   MEDLINE=92295566; PubMed=1318606;
RA   Telford E.A., Watson M.S., McBride K., Davison A.J.;
```

```
RT   "The DNA sequence of equine herpesvirus-1.";
RL   Virology 189:304-316(1992).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Ab4;
RA   Davison A.J.;
RL   Submitted (JUN-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY464052; AAS45909.1; -.
DR   EMBL; AY665713; AAT67282.1; -.
DR   GO; GO:0019028; C:viral capsid; IEA.
DR   InterPro; IPR007584; UL35.
DR   Pfam; PF04496; Herpes_UL35; 1.
SQ   SEQUENCE   119 AA;  13596 MW;  3603A1AA4208CB08 CRC64;

  Query Match            12.7%;  Score 153.5;  DB 2;  Length 119;
  Best Local Similarity   31.1%;  Pred. No. 0.039;
  Matches   37;  Conservative   16;  Mismatches   45;  Indels   21;  Gaps    3;

Qy          3 QPASSRVVFDPSNPTTFSVEAIAAYTPVALIRLLNASGPLQPGHRVDIADARSIYT---- 58
              |    : | |:|::|          || :| :||       :|  |    ||
Db          6 QQQQAPVAFNPADPPNIKAANFKDMLPVDVITILN--------QNIDELDYTK-YTEDEI 56

Qy         59 --------VGAAASAARARANHNANTIRRTAMFAETDPMTWLRPTVGLKRTFNPRIIRP 109
                    :| | :    |  |  : :||: |||: |  || || :|||||| || ::|
Db         57 SEGLKQLFMGTARTMVSLRQRHLKSLVRRSDMFAQNDASTWARPNIGLKRTFPPRFMQP 115


RESULT 12
Q66IV2
ID   Q66IV2      PRELIMINARY;      PRT;   503 AA.
AC   Q66IV2;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   MGC84356 protein.
GN   Name=MGC84356;
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Eye;
RX   MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA   Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA   Richardson P.;
RT   "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT   initiative.";
RL   Dev. Dyn. 225:384-391(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Eye;
RX   PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Eye;
RA   Klein S., Gerhard D.S.;
RL   Submitted (AUG-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Contains 1 SH3 domain.
DR   EMBL; BC081178; AAH81178.1; -.
DR   InterPro; IPR000108; Neu_cyt_fact_2.
DR   InterPro; IPR001452; SH3.
DR   InterPro; IPR011511; SH3_2.
DR   InterPro; IPR000727; T_SNARE.
DR   Pfam; PF00018; SH3_1; 1.
DR   Pfam; PF07653; SH3_2; 1.
DR   PRINTS; PR00499; P67PHOX.
DR   PRINTS; PR00452; SH3DOMAIN.
DR   ProDom; PD000066; SH3; 1.
DR   SMART; SM00326; SH3; 1.
DR   PROSITE; PS50002; SH3; 1.
DR   PROSITE; PS50192; T_SNARE; 1.
KW   SH3 domain.
SQ   SEQUENCE   503 AA;  54100 MW;  DED356AEE64A2C09 CRC64;

  Query Match            11.8%;  Score 142;  DB 2;  Length 503;
  Best Local Similarity   29.9%;  Pred. No. 0.79;
  Matches   53;  Conservative   20;  Mismatches   68;  Indels   36;  Gaps   10;

Qy         69 RANHNANTIRRTAMFAETDPMTWLRPTVGLKRTFNPRIIRPQPPNPSM-SLGISGP--TI 125
              :| |  |   ||     |:| |  :|              | || ||  :|| : |  |:
Db        157 KAKHGGNQAARTGTLTRTNPPT-QKP-------------PSPPMPSRGTLGRNTPYKTL 201

Qy        126 LPQKTQSADQSALQQPAALAFSGSSPQHPPPQTTSASVGQQQHVVSGSS-GQQPQQGAQS 184
               | |  :     :  || |   ||  || | :|  ||: |:    |||| |   :: : |
Db        202 EPVKPPTVPNDYMTSPARL---GS--QHSPGRT--ASLNQRPRTHSGSSGGSGSRENSGS 254

Qy        185 STV-----QPTTGSPPAAQGV-----PQSTPPPTQNTPQGGKGQTLSHTGQSGNASR 231
              |::      ||   |   ||:     |   |||    |     |   :   | |: :|
Db        255 SSIGIPIAVPTPSPPTVGQGISGGPPPPPAPPPPPPPPPPPVGTGSAPASQFGSMTR 311


RESULT 13
Q76N76
ID   Q76N76      PRELIMINARY;      PRT;   771 AA.
AC   Q76N76;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   GEF-1.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Wistar;
RX   MEDLINE=99013250; PubMed=9798906;
RA   Ogura K., Kohno K., Tai T.;
RT   "Molecular cloning of a rat brain cDNA, with homology to a tyrosine
RT   kinase substrate, that induces galactosylceramide expression in COS-7
RT   cells.";
RL   J. Neurochem. 71:1827-1836(1998).
DR   EMBL; AB002811; BAD08342.1; -.
DR   HSSP; Q9UJY5; 1JWG.
DR   GO; GO:0008270; F:zinc ion binding; IEA.
DR   GO; GO:0006886; P:intracellular protein transport; IEA.
DR   InterPro; IPR008942; ENTH_VHS.
DR   InterPro; IPR011011; FYVE_PHD_ZnF.
DR   InterPro; IPR003903; UIM.
DR   InterPro; IPR002014; VHS.
DR   InterPro; IPR000306; Znf_FYVE.
DR   Pfam; PF01363; FYVE; 1.
DR   Pfam; PF02809; UIM; 1.
DR   Pfam; PF00790; VHS; 1.
```

```
DR   ProDom; PD003686; VHS; 1.
DR   SMART; SM00064; FYVE; 1.
DR   SMART; SM00288; VHS; 1.
DR   PROSITE; PS50330; UIM; 1.
DR   PROSITE; PS50179; VHS; 1.
DR   PROSITE; PS50178; ZF_FYVE; 1.
SQ   SEQUENCE   771 AA;  85788 MW;  4A0DF0A058176739 CRC64;

  Query Match             11.7%;  Score 141.5;  DB 2;  Length 771;
  Best Local Similarity   28.2%;  Pred. No. 1.3;
  Matches   67;  Conservative   24;  Mismatches   86;  Indels   61;  Gaps   14;

Qy          4 PASSRVVFDPSNPTTF--------SVEAIAAYTPVALIRLLNASGPL--QPGHRVDIADA 53
              |  :  |:: || ||:|        |||    :  | : :   |:||    ||   |
Db        573 PTAGGVLYQPSGPTSFPGTFSPAGSVEGSPMH-GVYMSQPAPATGPYPSMPGTTAD-PSM 630

Qy         54 RSIYTVGAAASAARARANHNANTIRRTAMFAETDPMTWLRPTVGLKRTFNPRIIRPQP-P 112
               | |   | |  |:|             |  |:  | |             ||     || |
Db        631 VSAYMYPAGAPGAQA------------APQAQAGPTT------------NPAYSSYQPTP 666

Qy        113 NPSMSLGISGPTILPQKTQSADQSAL----QQPAALAFSGSSPQHPPPQTT----SASVG 164
              |    |   |  ||  :|    | :      || :: :   : |:    ||     ||:
Db        667 TP----GYQAPQSLPAISQPPQTSNIGYMGSQPMSMGYQPYNMQN--LMTTLPGQDASLP 720

Qy        165 QQQHVVSGSSGQQPQQGAQSSTVQPTTGSPPAAQGVPQSTPPPTQNTP-QGGKGQTLS 221
               ||  :   :||||     : :  |    || ||      |||||  | || : | :|
Db        721 AQQPYI---TGQQPMYQQMAPSTGPPQQQPPVAQ------PPPTQGPPAQGNETQLIS 769


RESULT 14
Q6P9L3
ID   Q6P9L3      PRELIMINARY;      PRT;  1114 AA.
AC   Q6P9L3;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Falz protein.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6; TISSUE=Brain;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6; TISSUE=Brain;
RA   Strausberg R.;
RL   Submitted (NOV-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC060715; AAH60715.1; -.
DR   GO; GO:0003677; F:DNA binding; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR001487; Bromodomain.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR001965; Znf_PHD.
DR   Pfam; PF00439; Bromodomain; 1.
DR   Pfam; PF00628; PHD; 1.
DR   PRINTS; PR00503; BROMODOMAIN.
DR   SMART; SM00297; BROMO; 1.
DR   SMART; SM00249; PHD; 1.
DR   PROSITE; PS00633; BROMODOMAIN_1; 1.
DR   PROSITE; PS50014; BROMODOMAIN_2; 1.
DR   PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR   PROSITE; PS01359; ZF_PHD_1; UNKNOWN_1.
DR   PROSITE; PS50016; ZF_PHD_2; 1.
SQ   SEQUENCE   1114 AA;  121422 MW;  7A3906159D59599B CRC64;

  Query Match             11.5%;  Score 139;  DB 2;  Length 1114;
  Best Local Similarity   27.0%;  Pred. No. 2.7;
  Matches   57;  Conservative   23;  Mismatches   77;  Indels   54;  Gaps    9;

Qy         54 RSIYT-VGAAASAARARANHNANTIRRTAMFA-----------ETDPMTWLRPTVGLKR 100
              | ::| :  :|:|| : :| :: |   ||  :           :  | | | ||
Db        413 RFLFTPLSTSATAASSSSNSSSTTTNATAAGSGEQKQSKILPQTQVQPATTLAPT--QSS 470

Qy        101 TFNPRIIRPQPPNPSMSLGISGPTILPQKTQSADQSALQQPAALAFSGSSPQHPPPQTTS 160
              : :|   :|||  |:     : |   ||    |          |||  :   |  |    ||:
Db        471 SVSPAEAQPQPAQPA-----AQPQ--PQPQPPAQPEVQTQPAVSSHVPSETQPSQAQTSK 523

Qy        161 ASVG---QQQHVVSGSS----------------------GQQPQQGAQSSTVQPTTGSPP 195
                |    | |  | | |                     |||||   |::  ||    ||
Db        524 PLVATQCQPQSSVQGQSPVRVQSPPLTRIRPSTPSQVTPGQQPQ--VQTTASQPIPIPPP 581

Qy        196 A-----AQGVPQSTPPPTQNTPQGGKGQTLS 221
                  :|| ||| |    :|     ||||:
Db        582 TSLQAPSQGQPQSQPQVQSSTQTLSSGQTLN 612


RESULT 15
Q9JJ50
ID   Q9JJ50      PRELIMINARY;      PRT;   776 AA.
AC   Q9JJ50;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hrs.
GN   Name=Hrs;
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20286537; PubMed=10825299;
RA   Kwong J., Roudabush F.L., Moore P.H., Montague M., Oldham W., Li Y.,
RA   Chin L.-S., Li L.;
RT   "Hrs interacts with SNAP-25 and regulates Ca2+-dependent exocytosis.";
RL   J. Cell Sci. 113:2273-2284(2000).
DR   EMBL; AF036344; AAF76251.1; -.
DR   HSSP; Q960X8; 1DVP.
DR   GO; GO:0008270; F:zinc ion binding; IEA.
DR   GO; GO:0006886; P:intracellular protein transport; IEA.
DR   InterPro; IPR008942; ENTH_VHS.
DR   InterPro; IPR011011; FYVE_PHD_ZnF.
DR   InterPro; IPR003903; UIM.
DR   InterPro; IPR002014; VHS.
DR   InterPro; IPR000306; Znf_FYVE.
DR   Pfam; PF01363; FYVE; 1.
DR   Pfam; PF02809; UIM; 1.
DR   Pfam; PF00790; VHS; 1.
DR   ProDom; PD003686; VHS; 1.
DR   SMART; SM00064; FYVE; 1.
DR   SMART; SM00726; UIM; 1.
DR   SMART; SM00288; VHS; 1.
DR   PROSITE; PS50330; UIM; 1.
```

```
DR   PROSITE; PS50179; VHS; 1.
DR   PROSITE; PS50178; ZF_FYVE; 1.
SQ   SEQUENCE   776 AA;  86246 MW;  FD61BEC121F37E67 CRC64;

  Query Match             11.4%;  Score 138;  DB 2;  Length 776;
  Best Local Similarity   27.6%;  Pred. No. 2.1;
  Matches   66;  Conservative   26;  Mismatches   89;  Indels   58;  Gaps   14;

Qy          4 PASSRVVFDPSNPTTF--------SVEAIAAYTPVALIRLLNASGPL--QPGHRVDIADA 53
              | :  |:: || ||:|        |||    :  | : :   |:||    ||   |
Db        573 PTAGGVLYQPSGPTSFPGTFSPAGSVEGSPMH-GVYMSQPAPATGPYPSMPGTTAD-PSM 630

Qy         54 RSIYTVGAAASAARARANHNANTIRRTAMFAETDPMTWLRPTVGLKRTFNPRIIRPQP-P 112
               | |   | |  |:|             |  |:  | |              ||     || |
Db        631 VSAYMYPAGAPGAQA-----------APQAQAGPTT-----------NPAYSSYQPTP 666

Qy        113 NPS-MSLGISGPTILPQKTQSADQSAL----QQPAALAFSGSSPQHPPPQTT----SASV 163
               |   ::    | || :|    | :     || :: :   : |:    ||     ||:
Db        667 TPGYQNVASQAPQSLPAISQPPQTSNIGYMGSQPMSMGYQPYNMQN--LMTTLPGQDASL 724

Qy        164 GQQQHVVSGSSGQQPQQGAQSSTVQPTTGSPPAAQGVPQSTPPPTQNTP-QGGKGQTLS 221
              ||  :   :||||     : :  |    || ||      |||||  | || : | :|
Db        725 PAQQPYI---TGQQPMYQQMAPSTGPPQQQPPVAQ------PPPTQGPPAQGNETQLIS 774


Search completed: September 25, 2005, 21:01:59
Job time : 391 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 25, 2005, 20:26:05 ; Search time 119 Seconds
(without alignments)
147.416 Million cell updates/sec

Title: US-09-874-140-2
Perfect score: 1206
Sequence: 1 MTQPASSRVVFDPSNPTTFS.....KGQTLSHTGQSGNASRSRRV 235

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1206	100.0	235	3	US-09-219-337-2	Sequence 2, Appli
2	128.5	10.7	480	3	US-09-189-035-5	Sequence 5, Appli
3	128.5	10.7	480	3	US-09-382-086-5	Sequence 5, Appli
4	127.5	10.6	531	4	US-09-949-016-8074	Sequence 8074, Ap
5	127.5	10.6	907	3	US-08-783-774-2	Sequence 2, Appli
6	127.5	10.6	907	4	US-09-328-599A-1	Sequence 1, Appli
7	127.5	10.6	907	5	PCT-US95-04611A-19	Sequence 19, Appl
8	123	10.2	276	4	US-09-538-092-889	Sequence 889, App
9	123	10.2	331	4	US-09-538-092-845	Sequence 845, App
10	123	10.2	529	4	US-09-949-016-11657	Sequence 11657, A
11	122	10.1	2441	1	US-08-194-468-2	Sequence 2, Appli
12	122	10.1	2441	3	US-08-961-739-2	Sequence 2, Appli
13	122	10.1	2441	3	US-09-514-247A-8	Sequence 8, Appli
14	122	10.1	2441	4	US-09-686-316-2	Sequence 2, Appli
15	121.5	10.1	258	4	US-09-949-016-9409	Sequence 9409, Ap
16	121	10.0	251	4	US-09-538-092-840	Sequence 840, App
17	120.5	10.0	234	4	US-09-538-092-888	Sequence 888, App
18	120	10.0	677	1	US-08-188-582-13	Sequence 13, Appl
19	120	10.0	677	1	US-08-646-715-13	Sequence 13, Appl
20	120	10.0	677	4	US-09-538-092-1164	Sequence 1164, Ap
21	120	10.0	694	4	US-09-949-016-8774	Sequence 8774, Ap
22	120	10.0	694	4	US-09-949-016-8775	Sequence 8775, Ap
23	120	10.0	711	4	US-09-949-016-8493	Sequence 8493, Ap
24	119	9.9	258	4	US-09-248-796A-23723	Sequence 23723, A
25	119	9.9	320	4	US-09-252-991A-21056	Sequence 21056, A
26	118.5	9.8	247	4	US-09-538-092-890	Sequence 890, App
27	118.5	9.8	542	4	US-09-538-092-289	Sequence 289, App
28	118	9.8	878	4	US-09-556-706B-2	Sequence 2, Appli
29	118	9.8	878	4	US-09-724-418A-2	Sequence 2, Appli
30	118	9.8	1109	4	US-09-949-016-10771	Sequence 10771, A
31	118	9.8	1203	4	US-09-949-016-6615	Sequence 6615, Ap
32	117	9.7	577	4	US-10-029-180-48	Sequence 48, Appl
33	117	9.7	2442	3	US-09-514-247A-10	Sequence 10, Appl
34	117	9.7	2442	4	US-09-538-092-1370	Sequence 1370, Ap
35	115.5	9.6	124	2	US-08-925-237-2	Sequence 2, Appli
36	115	9.5	419	4	US-09-248-796A-21670	Sequence 21670, A
37	115	9.5	666	4	US-09-050-739-70	Sequence 70, Appl
38	114.5	9.5	320	4	US-09-949-016-8196	Sequence 8196, Ap
39	114.5	9.5	593	4	US-09-252-991A-17983	Sequence 17983, A
40	113	9.4	662	4	US-09-902-540-12594	Sequence 12594, A
41	113	9.4	696	3	US-08-906-865-4	Sequence 4, Appli
42	113	9.4	696	4	US-09-129-668-4	Sequence 4, Appli
43	113	9.4	705	4	US-09-538-092-945	Sequence 945, App
44	113	9.4	713	4	US-09-949-016-9700	Sequence 9700, Ap
45	113	9.4	776	4	US-09-266-225D-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1
US-09-219-337-2
; Sequence 2, Application US/09219337
; Patent No. 6258363
; GENERAL INFORMATION:
; APPLICANT: EICKMANN, Markus
; APPLICANT: GICKLHORN, Dorothee
; APPLICANT: RADSAK, Klaus
; APPLICANT: HAUSER, Hans-Peter
; APPLICANT: GIESENDORF, Bernhard
; TITLE OF INVENTION: VARICELLA ZOSTER VIRUS (VZV) IMMUNOREACTIVE PROTEIN
; TITLE OF INVENTION: VP26 AND ITS DIAGNOSTIC USE
; FILE REFERENCE: 058315/0127
; CURRENT APPLICATION NUMBER: US/09/219,337
; CURRENT FILING DATE: 1998-12-23
; EARLIER APPLICATION NUMBER: DE 197 57 765.2
; EARLIER FILING DATE: 1997-12-23
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 235
; TYPE: PRT
; ORGANISM: Varicella Zoster Virus
US-09-219-337-2

Query Match 100.0%; Score 1206; DB 3; Length 235;
Best Local Similarity 100.0%; Pred. No. 1.3e-93;
Matches 235; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MTQPASSRVVFDPSNPTTFSVEAIAAYTPVALIRLLNASGPLQPGHRVDIADARSIYTVG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MTQPASSRVVFDPSNPTTFSVEAIAAYTPVALIRLLNASGPLQPGHRVDIADARSIYTVG 60

Qy         61 AAASAARARANHNANTIRRTAMFAETDPMTWLRPTVGLKRTFNPRIIRPQPPNPSMSLGI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AAASAARARANHNANTIRRTAMFAETDPMTWLRPTVGLKRTFNPRIIRPQPPNPSMSLGI 120

Qy        121 SGPTILPQKTQSADQSALQQPAALAFSGSSPQHPPPQTTSASVGQQQHVVSGSSGQQPQQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SGPTILPQKTQSADQSALQQPAALAFSGSSPQHPPPQTTSASVGQQQHVVSGSSGQQPQQ 180

Qy        181 GAQSSTVQPTTGSPPAAQGVPQSTPPPTQNTPQGGKGQTLSHTGQSGNASRSRRV 235
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GAQSSTVQPTTGSPPAAQGVPQSTPPPTQNTPQGGKGQTLSHTGQSGNASRSRRV 235
```

RESULT 2
US-09-189-035-5
; Sequence 5, Application US/09189035
; Patent No. 6020165

```
; GENERAL INFORMATION:
;  APPLICANT: Yue, Henry
;  APPLICANT: Corley, Neil C.
;  APPLICANT: Guegler, Karl J.
;  APPLICANT: Baughn, Mariah R.
;  TITLE OF INVENTION: CYTOKINE SIGNAL REGULATORS
;  FILE REFERENCE: PF-0638 US
;  CURRENT APPLICATION NUMBER: US/09/189,035
;  CURRENT FILING DATE: 1998-11-10
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: PERL Program
; SEQ ID NO 5
;   LENGTH: 480
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE: -
;   OTHER INFORMATION: g2245671
US-09-189-035-5

  Query Match             10.7%;  Score 128.5;  DB 3;  Length 480;
  Best Local Similarity   28.2%;  Pred. No. 0.0067;
  Matches   51;  Conservative   17;  Mismatches   66;  Indels   47;  Gaps    8;

Qy         69 RANHNANTIRRTAMFAETDPMTWLRPTVGLKRTFNPRIIRPQPPNPSMS----LGISGP- 123
              :| |  |   ||   : |:| |                   :||:| ||    || : |
Db        157 KAKHGNNQPARTGTLSRTNPPT-----------------QKPPSPPMSGRGTLGRNTPY 198

Qy        124 -TILPQKTQSADQSALQQPAALAFSGSSPQHPPPQTTSASVGQQQHV-VSGSSGQQPQQG 181
               |: | |  :     :  || |   ||  || | :| | :   : |   || || :   |
Db        199 KTLEPVKPPTVPNDYMTSPARL---GS--QHSPGRTASLNQRPRTHSGSSGGSGSRENSG 253

Qy        182 AQS--STVQPTTGSPPAAQGVPQSTP---------------PPTQNTPQGGKGQTLSHTG 224
              : |    :   | |||    | | |                 | :|  ||  :| | |
Db        254 SSSIGIPIAVPTPSPPTIGPAPGSAPGSQYGTMTRQISRHNSTTSSTSSGGYRRTPSVTA 313

Qy        225 Q 225
               |
Db        314 Q 314


RESULT 3
US-09-382-086-5
; Sequence 5, Application US/09382086
; Patent No. 6201106
; GENERAL INFORMATION:
;  APPLICANT: Yue, Henry
;  APPLICANT: Corley, Neil C.
;  APPLICANT: Guegler, Karl J.
;  APPLICANT: Baughn, Mariah R.
;  TITLE OF INVENTION: CYTOKINE SIGNAL REGULATORS
;  FILE REFERENCE: PF-0638 US
;  CURRENT APPLICATION NUMBER: US/09/382,086
;  CURRENT FILING DATE: 1999-08-24
;  EARLIER APPLICATION NUMBER: 09/189,035
;  EARLIER FILING DATE: 1998-11-10
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: PERL Program
; SEQ ID NO 5
;   LENGTH: 480
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE: -
;   OTHER INFORMATION: g2245671
US-09-382-086-5

  Query Match             10.7%;  Score 128.5;  DB 3;  Length 480;
  Best Local Similarity   28.2%;  Pred. No. 0.0067;
  Matches   51;  Conservative   17;  Mismatches   66;  Indels   47;  Gaps    8;

Qy         69 RANHNANTIRRTAMFAETDPMTWLRPTVGLKRTFNPRIIRPQPPNPSMS----LGISGP- 123
              :| |  |   ||   : |:| |                   :||:| ||    || : |
Db        157 KAKHGNNQPARTGTLSRTNPPT-----------------QKPPSPPMSGRGTLGRNTPY 198

Qy        124 -TILPQKTQSADQSALQQPAALAFSGSSPQHPPPQTTSASVGQQQHV-VSGSSGQQPQQG 181
               |: | |  :     :  || |   ||  || | :| | :   : |   || || :   |
Db        199 KTLEPVKPPTVPNDYMTSPARL---GS--QHSPGRTASLNQRPRTHSGSSGGSGSRENSG 253

Qy        182 AQS--STVQPTTGSPPAAQGVPQSTP---------------PPTQNTPQGGKGQTLSHTG 224
              : |    :   | |||    | | |                 | :|  ||  :| | |
Db        254 SSSIGIPIAVPTPSPPTIGPAPGSAPGSQYGTMTRQISRHNSTTSSTSSGGYRRTPSVTA 313

Qy        225 Q 225
               |
Db        314 Q 314


RESULT 4
US-09-949-016-8074
; Sequence 8074, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8074
;   LENGTH: 531
;   TYPE: PRT
;   ORGANISM: Human
US-09-949-016-8074

  Query Match             10.6%;  Score 127.5;  DB 4;  Length 531;
  Best Local Similarity   26.0%;  Pred. No. 0.0092;
  Matches   54;  Conservative   18;  Mismatches   57;  Indels   79;  Gaps    9;

Qy         72 HNANTIRRTAMFAETDPMTWLRPTVGLKRTFNPRIIRPQPPNPSMS----LGISGP--TI 125
              |  |   ||   : |:| |                   :||:| ||    || : |  |:
Db        183 HGNNQPARTGTLSRTNPPT-----------------QKPPSPPMSGRGTLGRNTPYKTL 224

Qy        126 LPQKTQSADQSALQQPAALAFSGSSPQHPPPQTTSASVGQQQHVVSGSSG---------- 175
               | |  :     :  || |   ||  || | :|  ||: |:    |||||
Db        225 EPVKPPTVPNDYMTSPARL---GS--QHSPGRT--ASLNQRPRTHSGSSGGSGSRENSGS 277

Qy        176 -----------QQPQQGAQSSTVQPTTGSPPAAQ-----------GVPQSTP------- 205
                        |  | :: :| | :|:|||              | | |
Db        278 SSIGIPIAVPTPSPPTIGPENISVPPPSGAPPAPPLAPLLPVSTVIAAPGSAPGSQYGTM 337

Qy        206 --------PPTQNTPQGGKGQTLSHTGQ 225
                      | :|  ||  :| | | |
Db        338 TRQISRHNSTTSSTSSGGYRRTPSVTAQ 365


RESULT 5
US-08-783-774-2
; Sequence 2, Application US/08783774
; Patent No. 6054130
;  GENERAL INFORMATION:
;    APPLICANT:  Spaete, Richard
;    APPLICANT:  Jackman, Winthrop
;    TITLE OF INVENTION:  NON-SPLICING VARIANTS OF
;    TITLE OF INVENTION:  GP350/220
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
```

```
;       STATE:  NY
;       COUNTRY:  USA
;       ZIP:  10036/2711
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Diskette
;       COMPUTER:  IBM Compatible
;       OPERATING SYSTEM:  DOS
;       SOFTWARE:  FastSEQ Version 2.0
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/783,774
;       FILING DATE:  15-JAN-1997
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Coruzzi, Laura A.
;       REGISTRATION NUMBER:  30,742
;       REFERENCE/DOCKET NUMBER:  7682-037
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  212-790-9090
;       TELEFAX:  212-869-8864
;       TELEX:  66141 PENNIE
;   INFORMATION FOR SEQ ID NO:  2:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  907 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:
;       TOPOLOGY:  unknown
;     MOLECULE TYPE:  protein
US-08-783-774-2

  Query Match             10.6%;  Score 127.5;  DB 3;  Length 907;
  Best Local Similarity   24.9%;  Pred. No. 0.018;
  Matches   64;  Conservative   26;  Mismatches   88;  Indels   79;  Gaps   14;

Qy          2 TQPASSRVVFDP--SNPT---TFSVEAIAAYTPVALIRLLNASGPLQPGHRVDIADARSI 56
              | | |:     |  ::||   |    |:   ||       ||:||
Db        577 TSPTSAVTTPTPNATSPTLGKTSPTSAVTTPTP-------NATGP-------------- 614

Qy         57 YTVGAAASAARARANHNANTIRRTAMFAETDPMTWLRPTVGLKRTFNPRIIRPQPPNPSM 116
               |||  :  | |  ||   |:  |:            ||          ::  || | :
Db        615 -TVGETSPQANA-TNH---TLGGTS-----------PT---------PVVTSQPKNATS 648

Qy        117 SLGISGPTILPQKTQSADQSALQQPAAL--AFSGSSPQHPP-------------PQTTSA 161
              ::      |   | |       |  | : | :|  | |             | | |
Db        649 AVTTGQHNITSSSTSSMSLRPSSNPETLSPSTSDNSTSHMPLLTSAHPTGGENITQVTPA 708

Qy        162 SVGQQQHVVSGSSGQQPQQGAQSSTVQP---TTGSPPAAQGVPQSTPPPTQNTPQGGKGQ 218
              |:    | || ||  :|: |  |    |   :| : |    | : |||    :||   ||
Db        709 SI--STHHVSTSS-PEPRPGTTSQASGPGNSSTSTKPGEVNVTKGTPPQNATSPQAPSGQ 765

Qy        219 -----TLSHTGQSGNAS 230
                  |:: ||   |::
Db        766 KTAVPTVTSTGGKANST 782


RESULT 6
US-09-328-599A-1
; Sequence 1, Application US/09328599A
; Patent No. 6432679
;  GENERAL INFORMATION:
;    APPLICANT:  MOND, James J. and Lees, Andrew
;    TITLE OF INVENTION:  Enhancement of B Cell Activation by
;    TITLE OF INVENTION:  Co-Ligation of Receptors for Antigen and Complement C3d
;    TITLE OF INVENTION:  Using EBV gp350/220 or EBV gp350/220 Peptide Adjuvants
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Finnegan, Henderson, Farabow, Garrett &
;      ADDRESSEE:  Dunner, L.L.P.
;      STREET:  1300 I Street, N.W., Suite 700
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/328,599A
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fordis, Jean B.
;      REGISTRATION NUMBER:  32,984
;      REFERENCE/DOCKET NUMBER:  04995.6025-00000
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)408-4000
;      TELEFAX:  (202)408-4400
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  907 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-328-599A-1

  Query Match             10.6%;  Score 127.5;  DB 4;  Length 907;
  Best Local Similarity   24.9%;  Pred. No. 0.018;
  Matches   64;  Conservative   26;  Mismatches   88;  Indels   79;  Gaps   14;

Qy          2 TQPASSRVVFDP--SNPT---TFSVEAIAAYTPVALIRLLNASGPLQPGHRVDIADARSI 56
              | | |:     |  ::||   |    |:   ||       ||:||
Db        577 TSPTSAVTTPTPNATSPTLGKTSPTSAVTTPTP-------NATGP-------------- 614

Qy         57 YTVGAAASAARARANHNANTIRRTAMFAETDPMTWLRPTVGLKRTFNPRIIRPQPPNPSM 116
               |||  :  | |  ||   |:  |:            ||          ::  || | :
Db        615 -TVGETSPQANA-TNH---TLGGTS-----------PT---------PVVTSQPKNATS 648

Qy        117 SLGISGPTILPQKTQSADQSALQQPAAL--AFSGSSPQHPP-------------PQTTSA 161
              ::      |   | |       |  | : | :|  | |             | | |
Db        649 AVTTGQHNITSSSTSSMSLRPSSNPETLSPSTSDNSTSHMPLLTSAHPTGGENITQVTPA 708

Qy        162 SVGQQQHVVSGSSGQQPQQGAQSSTVQP---TTGSPPAAQGVPQSTPPPTQNTPQGGKGQ 218
              |:    | || ||  :|: |  |    |   :| : |    | : |||    :||   ||
Db        709 SI--STHHVSTSS-PEPRPGTTSQASGPGNSSTSTKPGEVNVTKGTPPQNATSPQAPSGQ 765

Qy        219 -----TLSHTGQSGNAS 230
                  |:: ||   |::
Db        766 KTAVPTVTSTGGKANST 782


RESULT 7
PCT-US95-04611A-19
; Sequence 19, Application PC/TUS9504611A
;  GENERAL INFORMATION:
;    APPLICANT:  Spaete, Richard and Jackman, Winthrop, T.
;    TITLE OF INVENTION:  Non Splicing Variants of gp350/220
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Cooley Godward Castro Huddleson & Tatum
;      STREET:  5 Palo Alto Square
;      CITY:  Palo Alto
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94306
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  Patent In Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/04611A
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
```

```
;      APPLICATION NUMBER:  08/229,291
;      FILING DATE:  April 18, 1994
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Luann Cserr
;      REGISTRATION NUMBER:  31,822
;      REFERENCE/DOCKET NUMBER:  AVIR-003/00US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-843-5163
;      TELEFAX:  415-857-0663
;      TELEX:  380816 CooleyPA
;  INFORMATION FOR SEQ ID NO:  19:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  907 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
PCT-US95-04611A-19

  Query Match             10.6%;  Score 127.5;  DB 5;  Length 907;
  Best Local Similarity   24.9%;  Pred. No. 0.018;
  Matches   64;  Conservative   26;  Mismatches   88;  Indels   79;  Gaps   14;

Qy          2 TQPASSRVVFDP--SNPT---TFSVEAIAAYTPVALIRLLNASGPLQPGHRVDIADARSI 56
              | | |:     |  ::||   |    |:   ||       ||:||
Db        577 TSPTSAVTTPTPNATSPTLGKTSPTSAVTTPTP-------NATGP--------------- 614

Qy         57 YTVGAAASAARARANHNANTIRRTAMFAETDPMTWLRPTVGLKRTFNPRIIRPQPPNPSM 116
               |||  :  | |  ||   |:  |:            ||          ::  || | :
Db        615 -TVGETSPQANA-TNH---TLGGTS-----------PT---------PVVTSQPKNATS 648

Qy        117 SLGISGPTILPQKTQSADQSALQQPAAL--AFSGSSPQHPP-------------PQTTSA 161
              ::      |   | ||        |  |  : | :|  | |               | | |
Db        649 AVTTGQHNITSSSTSSMSLRPSSNPETLSPSTSDNSTSHMPLLTSAHPTGGENITQVTPA 708

Qy        162 SVGQQQHVVSGSSGQQPQQGAQSSTVQP---TTGSPPAAQGVPQSTPPPTQNTPQGGKGQ 218
              |:    | || ||  :|: |  |    |   :| : |    | : |||    :||   ||
Db        709 SI--STHHVSTSS-PEPRPGTTSQASGPGNSSTSTKPGEVNVTKGTPPQNATSPQAPSGQ 765

Qy        219 -----TLSHTGQSGNAS 230
                  |:: ||   |::
Db        766 KTAVPTVTSTGGKANST 782


RESULT 8
US-09-538-092-889
; Sequence 889, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
;  APPLICANT: Giot, Loic
;  APPLICANT:  Mansfield, Traci A.
;  TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
;  FILE REFERENCE: 15966-542
;  CURRENT APPLICATION NUMBER: US/09/538,092
;  CURRENT FILING DATE:  2000-03-29
;  PRIOR APPLICATION NUMBER: 60/127,352
;  PRIOR FILING DATE: 1999-04-01
;  PRIOR APPLICATION NUMBER: 60/178,965
;  PRIOR FILING DATE: 2000-02-01
;  NUMBER OF SEQ ID NOS: 1387
;  SOFTWARE: CuraPatSeqFormatter Version 0.9
; SEQ ID NO 889
;   LENGTH: 276
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (0)...(0)
;   OTHER INFORMATION: Polypeptide Accession Number P10162
US-09-538-092-889

  Query Match             10.2%;  Score 123;  DB 4;  Length 276;
  Best Local Similarity   32.1%;  Pred. No. 0.0099;
  Matches   45;  Conservative    6;  Mismatches   71;  Indels   18;  Gaps    5;

Qy        108 RPQPPNPSMSLGISGPTILPQKTQSADQSALQQPAALAFSGSSPQHPPPQTTSASVGQQQ 167
              :|| | |        ||   | | :          |              |: ||||  : | |
Db         22 KPQGPPPQGGNQSQGPPPPPGKPEGRPPQGGNQSQGPPPHPGKPERPPPQGGNQSQGPPP 81

Qy        168 HVVSGSSGQQPQQGAQSSTVQPTTGSP--PAAQGVPQS--TPP----PTQNTPQGGK--- 216
              |      | : || | ||    || | |  |  ||  ||  |||     |    ||||
Db         82 HPGKPES-RPPQGGHQSQGPPPTPGKPEGPPPQGGNQSQGTPPPPGKPEGRPPQGGNQSQ 140

Qy        217 ------GQTLSHTGQSGNAS 230
                   |:      | || |
Db        141 GPPPHPGKPERPPPQGGNQS 160


RESULT 9
US-09-538-092-845
; Sequence 845, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
;  APPLICANT: Giot, Loic
;  APPLICANT:  Mansfield, Traci A.
;  TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
;  FILE REFERENCE: 15966-542
;  CURRENT APPLICATION NUMBER: US/09/538,092
;  CURRENT FILING DATE:  2000-03-29
;  PRIOR APPLICATION NUMBER: 60/127,352
;  PRIOR FILING DATE: 1999-04-01
;  PRIOR APPLICATION NUMBER: 60/178,965
;  PRIOR FILING DATE: 2000-02-01
;  NUMBER OF SEQ ID NOS: 1387
;  SOFTWARE: CuraPatSeqFormatter Version 0.9
; SEQ ID NO 845
;   LENGTH: 331
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (0)...(0)
;   OTHER INFORMATION: Polypeptide Accession Number P04280
US-09-538-092-845

  Query Match             10.2%;  Score 123;  DB 4;  Length 331;
  Best Local Similarity   32.4%;  Pred. No. 0.012;
  Matches   44;  Conservative    7;  Mismatches   59;  Indels   26;  Gaps    6;

Qy        108 RPQPPNPSMSLGISGPTILPQKTQSADQSALQQPAALAFSGSSPQHPPPQ-----TTSAS 162
              :|| | |        ||   | | |    |     :|       |  || ||||     :  :
Db        157 KPQGPPPQGGNQPQGPPPPPGKPQGPPQQGGNRPQGPPPPG-KPQGPPPQGDKSRSPQSP 215

Qy        163 VGQQQHVVSGSSGQQPQQGAQSSTVQPTTGSPPAAQGVPQSTPPPTQNTPQG----GKGQ 218
              |: |        |  || | |              ||   | ||  ||   | |||    || |
Db        216 PGKPQ-------GPPPQGGNQPQ-------GPPPPPGKPQGPPPQGGNKPQGPPPPGKPQ 261

Qy        219 TLSHTGQSGNASRSRR 234
                  | |: |:| |
Db        262 --GPPAQGGSKSQSAR 275


RESULT 10
US-09-949-016-11657
; Sequence 11657, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
```

```
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11657
;   LENGTH: 529
;   TYPE: PRT
;   ORGANISM: Human
US-09-949-016-11657

  Query Match            10.2%;  Score 123;  DB 4;  Length 529;
  Best Local Similarity   27.9%;  Pred. No. 0.022;
  Matches   60;  Conservative   24;  Mismatches  105;  Indels   26;  Gaps    9;

Qy         13 PSNPTTFSVEAIAAYTPVALIRLLNASGPLQPGHRVDIADARSIYTVGAAASAARARANH 72
              |: |: ||        ||:   | |      |:    | : |   :  |||   | | |
Db        227 PTAPSVFS--PTGNRTPIPPSRTL-----LRKERGVKLLDISELDMVGAGREAKRRRKTL 279

Qy         73 NANTIRRTAMFAETDPMTWLRP--TVGLKRTFNPRIIRPQPPNPSMSLGISGPTILPQKT 130
              :|  : : |   |   :    |    ||  |     : :|  || |   | |:::|
Db        280 DAEVVEKPA--KEETVVENATPDYAAGLVSTQKLGSLNNEPALPSTSYLPSTPSVVP--A 335

Qy        131 QSADQSALQQPAALAFSGSSPQHPP--PQTTSASVGQQQHVV--SGSSGQQPQQGAQSS- 185
              |   |:   ||  :   | |   |  | | |  :  :|:  :  || |    |   | :|
Db        336 SSYIPSSETPPAPSSREASRPPEEPSAPSPTLPAQFKQRAPMYNSGLSPATPTPAAPTSP 395

Qy        186 -------TVQPTTGSPPAAQGVPQSTPPPTQNTPQ 213
                    | ||| :|| |   || | | |  |:
Db        396 LTPTTPPAVAPTTQTPPVAMVAPQ-TQAPAQQQPK 429


RESULT 11
US-08-194-468-2
; Sequence 2, Application US/08194468
; Patent No. 5750336
;  GENERAL INFORMATION:
;    APPLICANT:  Montminy, Marc R.
;    TITLE OF INVENTION:  ASSAYS FOR THE IDENTIFICATION OF
;    TITLE OF INVENTION:  COMPOUNDS WHICH INHIBIT ACTIVATION OF cAMP AND MITOGEN
;    TITLE OF INVENTION:  RESPONSIVE GENES
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pretty, Schroeder, Brueggemann & Clark
;      STREET:  444 South Flower Street, Suite 2000
;      CITY:  Los Angeles
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  90071
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/194,468
;      FILING DATE:  10-FEB-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Reiter, Stephen E.
;      REGISTRATION NUMBER:  31,192
;      REFERENCE/DOCKET NUMBER:  P41 9672
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619)-546-4737
;      TELEFAX:  (619)-546-9392
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2441 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-194-468-2

  Query Match            10.1%;  Score 122;  DB 1;  Length 2441;
  Best Local Similarity   26.8%;  Pred. No. 0.17;
  Matches   81;  Conservative   24;  Mismatches  103;  Indels   94;  Gaps   18;

Qy          4 PASSRVVFDPSNPTTFSV------EAIAAYTPVALIRLLNASGPLQP------GHRVDIA 51
              || |  |  |: |    |      : : :: |::|  :     |: |       | | :
Db        696 PAQS--VRPPNGPLPLPVNRMQVSQGMNSFNPMSLGNVQLPQAPMGPRAASPMNHSVQMN 753

Qy         52 DARSIYTVGAAASAARARANHN-----ANTIRRTAMFAETDPMTWLRPTVGLKRTFNPRI 106
                |:   | | | :|      |     || |   |             ||
Db        754 SMASV--PGMAISPSRMPQPPNMMGTHANNIMAQA-----------PT--------QNQ 791

Qy        107 IRPQPPNPSMSLGISGPTI-LPQKTQSADQSALQQPAA--------LAFSGS-------- 149
                ||   || |  :|  :: : |    |  |  |:| |         ||   |
Db        792 FLPQNQFPSSSGAMSVNSVGMGQPAAQAGVSQGQEPGAALPNPLNMLAPQASQLPCPPVT 851

Qy        150 -SPQH--PPPQTTSASVGQQQH-VVSGSSGQQPQQGAQSS-------TVQPTTGSPPAA- 197
               || |  ||| :|:| :   ||    | :  ||    | |       |  || || |:|
Db        852 QSPLHPTPPPASTAAGMPSLQHPTAPGMTPPQPAAPTQPSTPVSSGQTPTPTPGSVPSAA 911

Qy        198 --QGVP--------QSTP-------PPTQNTPQGGKGQ-TLSHTGQSGN------ASRSR 233
                | |         | ||        ||:  |||  : | |  ||   |       ||
Db        912 QTQSTPTVQAAAQAQVTPQPQTPVQPPSVATPQSSQQQPTPVHTQPPGTPLSQAAASIDN 971

Qy        234 RV 235
              ||
Db        972 RV 973


RESULT 12
US-08-961-739-2
; Sequence 2, Application US/08961739A
; Patent No. 6063583
; GENERAL INFORMATION:
;  APPLICANT: Montminy, Marc R.
;  TITLE OF INVENTION: Methods for Treating Diabetes Mellitus
;  FILE REFERENCE: SALK1650-1
;  CURRENT APPLICATION NUMBER: US/08/961,739A
;  CURRENT FILING DATE: 1997-10-31
;  EARLIER APPLICATION NUMBER: US 194,468
;  EARLIER FILING DATE: 1994-02-10
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 2441
;   TYPE: PRT
;   ORGANISM: Mus
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: (1)...(2441)
;   OTHER INFORMATION: Xaa = Any Amino Acid
US-08-961-739-2

  Query Match            10.1%;  Score 122;  DB 3;  Length 2441;
  Best Local Similarity   26.8%;  Pred. No. 0.17;
  Matches   81;  Conservative   24;  Mismatches  103;  Indels   94;  Gaps   18;

Qy          4 PASSRVVFDPSNPTTFSV------EAIAAYTPVALIRLLNASGPLQP------GHRVDIA 51
              || |  |  |: |    |      : : :: |::|  :     |: |       | | :
Db        696 PAQS--VRPPNGPLPLPVNRMQVSQGMNSFNPMSLGNVQLPQAPMGPRAASPMNHSVQMN 753

Qy         52 DARSIYTVGAAASAARARANHN-----ANTIRRTAMFAETDPMTWLRPTVGLKRTFNPRI 106
                |:   | | | :|      |     || |   |             ||
Db        754 SMASV--PGMAISPSRMPQPPNMMGTHANNIMAQA-----------PT--------QNQ 791

Qy        107 IRPQPPNPSMSLGISGPTI-LPQKTQSADQSALQQPAA--------LAFSGS-------- 149
                ||   || |  :|  :: : |    |  |  |:| |         ||   |
Db        792 FLPQNQFPSSSGAMSVNSVGMGQPAAQAGVSQGQEPGAALPNPLNMLAPQASQLPCPPVT 851

Qy        150 -SPQH--PPPQTTSASVGQQQH-VVSGSSGQQPQQGAQSS-------TVQPTTGSPPAA- 197
               || |  ||| :|:| :   ||    | :  ||    | |       |  || || |:|
Db        852 QSPLHPTPPPASTAAGMPSLQHPTAPGMTPPQPAAPTQPSTPVSSGQTPTPTPGSVPSAA 911
```

```
Qy        198 --QGVP--------QSTP-------PPTQNTPQGGKGQ-TLSHTGQSGN------ASRSR 233
                |  |        | ||       ||:  |||  : | |  ||   |       ||
Db        912 QTQSTPTVQAAAQAQVTPQPQTPVQPPSVATPQSSQQQPTPVHTQPPGTPLSQAAASIDN 971

Qy        234 RV 235
              ||
Db        972 RV 973


RESULT 13
US-09-514-247A-8
; Sequence 8, Application US/09514247A
; Patent No. 6365361
; GENERAL INFORMATION:
;  APPLICANT: TANABE SEIYAKU CO. LTD.
;  APPLICANT:  TANIGUCHI, Tomoyasu
;  APPLICANT:  MIZUKAMI, Junko
;  TITLE OF INVENTION: METHOD FOR IDENTIFYING OR SCREENING AGONIST AND ANTAGONIST TO PPA
;  FILE REFERENCE: TANIGUCHI=6
;  CURRENT APPLICATION NUMBER: US/09/514,247A
;  CURRENT FILING DATE:  2000-02-28
;  PRIOR APPLICATION NUMBER: PCT/JP98/03734
;  PRIOR FILING DATE: 1998-08-24
;  PRIOR APPLICATION NUMBER: JP231084/1997
;  PRIOR FILING DATE: 1997-08-27
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
;   LENGTH: 2441
;   TYPE: PRT
;   ORGANISM: mouse
US-09-514-247A-8

  Query Match            10.1%;  Score 122;  DB 3;  Length 2441;
  Best Local Similarity   26.8%;  Pred. No. 0.17;
  Matches   81;  Conservative   24;  Mismatches  103;  Indels   94;  Gaps   18;

Qy          4 PASSRVVFDPSNPTTFSV------EAIAAYTPVALIRLLNASGPLQP------GHRVDIA 51
              || |  |  |: |    |       : : :: |::|  :     |: |       | | :
Db        696 PAQS--VRPPNGPLPLPVNRMQVSQGMNSFNPMSLGNVQLPQAPMGPRAASPMNHSVQMN 753

Qy         52 DARSIYTVGAAASAARARANHN-----ANTIRRTAMFAETDPMTWLRPTVGLKRTFNPRI 106
                |:   | | | :|      |     || |   |            ||
Db        754 SMASV--PGMAISPSRMPQPPNMMGTHANNIMAQA-----------PT--------QNQ 791

Qy        107 IRPQPPNPSMSLGISGPTI-LPQKTQSADQSALQQPAA--------LAFSGS-------- 149
               ||    || |  :|  :: : |     |  |  |:| |        ||   |
Db        792 FLPQNQFPSSSGAMSVNSVGMGQPAAQAGVSQGQEPGAALPNPLNMLAPQASQLPCPPVT 851

Qy        150 -SPQH--PPPQTTSASVGQQQH-VVSGSSGQQPQQGAQSS-------TVQPTTGSPPAA- 197
                || |  ||| :|:| :   ||     | :  ||     | |       |  || || |:|
Db        852 QSPLHPTPPPASTAAGMPSLQHPTAPGMTPPQPAAPTQPSTPVSSGQTPTPTPGSVPSAA 911

Qy        198 --QGVP--------QSTP-------PPTQNTPQGGKGQ-TLSHTGQSGN------ASRSR 233
                |  |        | ||       ||:  |||  : | |  ||   |       ||
Db        912 QTQSTPTVQAAAQAQVTPQPQTPVQPPSVATPQSSQQQPTPVHTQPPGTPLSQAAASIDN 971

Qy        234 RV 235
              ||
Db        972 RV 973


RESULT 14
US-09-686-316-2
; Sequence 2, Application US/09686316
; Patent No. 6646115
; GENERAL INFORMATION:
;  APPLICANT: Montminy, Marc R.
;  TITLE OF INVENTION: Methods for Treating Diabetes Mellitus
;  FILE REFERENCE: SALK1650-1
;  CURRENT APPLICATION NUMBER: US/09/686,316
;  CURRENT FILING DATE:  2000-10-10
;  PRIOR APPLICATION NUMBER: US/08/961,739
;  PRIOR FILING DATE: 1997-10-31
;  PRIOR APPLICATION NUMBER: US 194,468
;  PRIOR FILING DATE: 1994-02-10
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 2441
;   TYPE: PRT
;   ORGANISM: Mus
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: (1)...(2441)
;   OTHER INFORMATION: Xaa = Any Amino Acid
US-09-686-316-2

  Query Match            10.1%;  Score 122;  DB 4;  Length 2441;
  Best Local Similarity   26.8%;  Pred. No. 0.17;
  Matches   81;  Conservative   24;  Mismatches  103;  Indels   94;  Gaps   18;

Qy          4 PASSRVVFDPSNPTTFSV------EAIAAYTPVALIRLLNASGPLQP------GHRVDIA 51
              || |  |  |: |    |       : : :: |::|  :     |: |       | | :
Db        696 PAQS--VRPPNGPLPLPVNRMQVSQGMNSFNPMSLGNVQLPQAPMGPRAASPMNHSVQMN 753

Qy         52 DARSIYTVGAAASAARARANHN-----ANTIRRTAMFAETDPMTWLRPTVGLKRTFNPRI 106
                |:   | | | :|      |     || |   |            ||
Db        754 SMASV--PGMAISPSRMPQPPNMMGTHANNIMAQA-----------PT--------QNQ 791

Qy        107 IRPQPPNPSMSLGISGPTI-LPQKTQSADQSALQQPAA--------LAFSGS-------- 149
               ||    || |  :|  :: : |     |  |  |:| |        ||   |
Db        792 FLPQNQFPSSSGAMSVNSVGMGQPAAQAGVSQGQEPGAALPNPLNMLAPQASQLPCPPVT 851

Qy        150 -SPQH--PPPQTTSASVGQQQH-VVSGSSGQQPQQGAQSS-------TVQPTTGSPPAA- 197
                || |  ||| :|:| :   ||     | :  ||     | |       |  || || |:|
Db        852 QSPLHPTPPPASTAAGMPSLQHPTAPGMTPPQPAAPTQPSTPVSSGQTPTPTPGSVPSAA 911

Qy        198 --QGVP--------QSTP-------PPTQNTPQGGKGQ-TLSHTGQSGN------ASRSR 233
                |  |        | ||       ||:  |||  : | |  ||   |       ||
Db        912 QTQSTPTVQAAAQAQVTPQPQTPVQPPSVATPQSSQQQPTPVHTQPPGTPLSQAAASIDN 971

Qy        234 RV 235
              ||
Db        972 RV 973


RESULT 15
US-09-949-016-9409
; Sequence 9409, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 9409
;   LENGTH: 258
;   TYPE: PRT
;   ORGANISM: Human
US-09-949-016-9409

  Query Match            10.1%;  Score 121.5;  DB 4;  Length 258;
  Best Local Similarity   24.8%;  Pred. No. 0.012;
  Matches   55;  Conservative   16;  Mismatches  102;  Indels   49;  Gaps    8;
```

```
Qy          20 SVEAIAAYTPVALIRLLNASGPLQPGHRVDIADARSIYTVGAAASAARARANHNANTIRR 79
               | | :      |||: | :|    :     |::   |:: :        | :  :
Db           9 SSEMLLILLSVALLALSSAESSSE-----DVSQEESLFLISGKPEGRRPQGGNQPQ---- 59

Qy          80 TAMFAETDPMTWLRPTVGLKRTFNPRIIRPQPPNPSMSLGISGPTILPQKTQSADQSALQ 139
                           ||         |   :|| | |       ||   | | :
Db          60 -------------RPP--------PPPGKPQGPPPQGGNQSQGPPPPPGKPEGRPPQGGN 98

Qy         140 QPAALAFSGSSPQHPPPQTTSASVGQQQHVVSGSSGQ----QPQQGAQSSTVQPTTGSP- 194
               |           |: ||||      | |         |:     || | ||    |  | |
Db          99 QSQGPPPHPGKPERPPPQG-----GNQSQGTPPPPGKPERPPPQGGNQSHRPPPPPGKPE 153

Qy         195 -PAAQGVPQST-PPPTQNTPQGGKGQTLSHTGQSGNASRSRR 234
                 |  ||  ||  |||    |:|          | || ||| |
Db         154 RPPPQGGNQSQGPPPHPGKPEG-------PPPQEGNKSRSAR 188
```

Search completed: September 25, 2005, 21:05:26
Job time : 120 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 25, 2005, 21:02:06 ; Search time 508 Seconds
(without alignments)
188.283 Million cell updates/sec

Title: US-09-874-140-2
Perfect score: 1206
Sequence: 1 MTQPASSRVVFDPSNPTTFS.....KGQTLSHTGQSGNASRSRRV 235

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1826521 seqs, 407012169 residues

Total number of hits satisfying chosen parameters: 1826521

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
19: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep:*
20: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
21: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1206	100.0	235	12	US-09-874-140-2	Sequence 2, Appli
2	153.5	12.7	119	17	US-10-626-832-31	Sequence 31, Appl
3	146	12.1	1367	15	US-10-320-797-3355	Sequence 3355, Ap
4	131	10.9	574	16	US-10-437-963-166255	Sequence 166255,
5	130	10.8	717	15	US-10-425-114-58581	Sequence 58581, A
6	128	10.6	1151	10	US-09-825-751A-79	Sequence 79, Appl
7	128	10.6	1151	18	US-10-851-438-79	Sequence 79, Appl
8	127.5	10.6	797	14	US-10-156-761-10907	Sequence 10907, A
9	127.5	10.6	907	9	US-09-328-599A-1	Sequence 1, Appli
10	127	10.5	630	16	US-10-322-281-428	Sequence 428, App
11	127	10.5	720	16	US-10-322-281-432	Sequence 432, App
12	126.5	10.5	839	14	US-10-245-752-54	Sequence 54, Appl
13	126.5	10.5	839	14	US-10-245-859-54	Sequence 54, Appl
14	126.5	10.5	839	14	US-10-245-103-54	Sequence 54, Appl
15	126.5	10.5	839	14	US-10-245-107-54	Sequence 54, Appl
16	126.5	10.5	839	14	US-10-245-143-54	Sequence 54, Appl
17	126.5	10.5	839	14	US-10-245-771-54	Sequence 54, Appl
18	126.5	10.5	839	14	US-10-245-851-54	Sequence 54, Appl
19	126.5	10.5	839	14	US-10-245-883-54	Sequence 54, Appl
20	126.5	10.5	839	14	US-10-237-535-54	Sequence 54, Appl
21	126.5	10.5	839	14	US-10-238-183-54	Sequence 54, Appl
22	126.5	10.5	839	14	US-10-238-283-54	Sequence 54, Appl
23	126.5	10.5	839	14	US-10-238-370-54	Sequence 54, Appl
24	126.5	10.5	839	14	US-10-245-055-54	Sequence 54, Appl
25	126.5	10.5	839	14	US-10-245-147-54	Sequence 54, Appl
26	126.5	10.5	839	14	US-10-245-730-54	Sequence 54, Appl
27	126.5	10.5	839	14	US-10-245-739-54	Sequence 54, Appl
28	126.5	10.5	839	14	US-10-246-210-54	Sequence 54, Appl
29	126.5	10.5	839	14	US-10-239-196-54	Sequence 54, Appl
30	126.5	10.5	839	14	US-10-243-024-54	Sequence 54, Appl
31	126.5	10.5	839	14	US-10-243-409-54	Sequence 54, Appl
32	126.5	10.5	839	14	US-10-245-621-54	Sequence 54, Appl
33	126.5	10.5	839	14	US-10-245-880-54	Sequence 54, Appl
34	126.5	10.5	839	14	US-10-245-033-54	Sequence 54, Appl
35	126.5	10.5	839	14	US-10-243-095-54	Sequence 54, Appl
36	126.5	10.5	839	14	US-10-245-185-54	Sequence 54, Appl
37	126.5	10.5	839	14	US-10-245-427-54	Sequence 54, Appl
38	126.5	10.5	839	14	US-10-245-473-54	Sequence 54, Appl
39	126.5	10.5	839	14	US-10-245-770-54	Sequence 54, Appl
40	126.5	10.5	839	14	US-10-245-877-54	Sequence 54, Appl
41	126.5	10.5	839	14	US-10-246-976-54	Sequence 54, Appl
42	126.5	10.5	839	14	US-10-243-320-54	Sequence 54, Appl
43	126.5	10.5	839	14	US-10-242-743-54	Sequence 54, Appl
44	126.5	10.5	839	14	US-10-242-845-54	Sequence 54, Appl
45	126.5	10.5	839	14	US-10-237-636-54	Sequence 54, Appl

ALIGNMENTS

RESULT 1
US-09-874-140-2
; Sequence 2, Application US/09874140
; Publication No. US20050202413A1
; GENERAL INFORMATION:
; APPLICANT: EICKMANN, Markus
; APPLICANT: GICKLHORN, Dorothee
; APPLICANT: RADSAK, Klaus
; APPLICANT: HAUSER, Hans-Peter
; APPLICANT: GIESENDORF, Bernhard
; TITLE OF INVENTION: VARICELLA ZOSTER VIRUS (VZV) IMMUNOREACTIVE PROTEIN
; TITLE OF INVENTION: VP26 AND ITS DIAGNOSTIC USE
; FILE REFERENCE: 058315/0127
; CURRENT APPLICATION NUMBER: US/09/874,140
; CURRENT FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US/09/219,337
; PRIOR FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: DE 197 57 765.2
; PRIOR FILING DATE: 1997-12-23
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 235
; TYPE: PRT
; ORGANISM: Varicella Zoster Virus
US-09-874-140-2

Query Match 100.0%; Score 1206; DB 12; Length 235;
Best Local Similarity 100.0%; Pred. No. 1.1e-80;
Matches 235; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MTQPASSRVVFDPSNPTTFSVEAIAAYTPVALIRLLNASGPLQPGHRVDIADARSIYTVG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MTQPASSRVVFDPSNPTTFSVEAIAAYTPVALIRLLNASGPLQPGHRVDIADARSIYTVG 60
```

```
Qy          61 AAASAARARANHNANTIRRTAMFAETDPMTWLRPTVGLKRTFNPRIIRPQPPNPSMSLGI 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 AAASAARARANHNANTIRRTAMFAETDPMTWLRPTVGLKRTFNPRIIRPQPPNPSMSLGI 120

Qy         121 SGPTILPQKTQSADQSALQQPAALAFSGSSPQHPPPQTTSASVGQQQHVVSGSSGQQPQQ 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 SGPTILPQKTQSADQSALQQPAALAFSGSSPQHPPPQTTSASVGQQQHVVSGSSGQQPQQ 180

Qy         181 GAQSSTVQPTTGSPPAAQGVPQSTPPPTQNTPQGGKGQTLSHTGQSGNASRSRRV 235
               |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 GAQSSTVQPTTGSPPAAQGVPQSTPPPTQNTPQGGKGQTLSHTGQSGNASRSRRV 235
```

RESULT 2
US-10-626-832-31
; Sequence 31, Application US/10626832
; Publication No. US20050003342A1
; GENERAL INFORMATION:
; APPLICANT: Davis Poynter, Nick
; APPLICANT: Nugent, Josephine
; APPLICANT: Birch-Machin, Ian
; APPLICANT: Allen, George P
; TITLE OF INVENTION: Viral Marker
; FILE REFERENCE: 620-262
; CURRENT APPLICATION NUMBER: US/10/626,832
; CURRENT FILING DATE: 2003-07-25
; PRIOR APPLICATION NUMBER: US 60/398,576
; PRIOR FILING DATE: 2002-07-26
; NUMBER OF SEQ ID NOS: 259
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 31
; LENGTH: 119
; TYPE: PRT
; ORGANISM: Equine herpesvirus 1
US-10-626-832-31

Query Match 12.7%; Score 153.5; DB 17; Length 119;
Best Local Similarity 31.1%; Pred. No. 0.00086;
Matches 37; Conservative 16; Mismatches 45; Indels 21; Gaps 3;

```
Qy           3 QPASSRVVFDPSNPTTFSVEAIAAYTPVALIRLLNASGPLQPGHRVDIADARSIYT---- 58
               |   : | |:|::|              || :| :||          :|  |    ||
Db           6 QQQQAPVAFNPADPPNIKAANFKDMLPVDVITILN--------QNIDELDYTK-YTEDEI 56

Qy          59 --------VGAAASAARARANHNANTIRRTAMFAETDPMTWLRPTVGLKRTFNPRIIRP 109
                       :| | :    | |  : :||: |||: |  || || :|||||| || ::|
Db          57 SEGLKQLFMGTARTMVSLRQRHLKSLVRRSDMFAQNDASTWARPNIGLKRTFPPRFMQP 115
```

RESULT 3
US-10-320-797-3355
; Sequence 3355, Application US/10320797
; Publication No. US20040014955A1
; GENERAL INFORMATION:
; APPLICANT: Eroshkin, Alexey M.
; APPLICANT: Zamudio, Carlos
; TITLE OF INVENTION: IDENTIFICATION OF ESSENTIAL GENES OF CRYPTOCOCCUS NEOFORMANS AND
; TITLE OF INVENTION: METHODS OF USE
; FILE REFERENCE: 10182-021-999
; CURRENT APPLICATION NUMBER: US/10/320,797
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: 60/341,261
; PRIOR FILING DATE: 2001-12-17
; NUMBER OF SEQ ID NOS: 3361
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3355
; LENGTH: 1367
; TYPE: PRT
; ORGANISM: Cryptococcus neoformans
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (357)..(357)
; OTHER INFORMATION: Xaa = any amino acid
US-10-320-797-3355

Query Match 12.1%; Score 146; DB 15; Length 1367;
Best Local Similarity 25.5%; Pred. No. 0.046;
Matches 70; Conservative 25; Mismatches 68; Indels 112; Gaps 14;

```
Qy           5 ASSRVVFDPSNPTTFSVEAIAAY-TPVALIRLLNASGPLQPGHRVDIADARSIYTVGAAA 63
               |||:  :    |||::     ||  | |      |||| ||                 |:
Db         944 ASSQPAYALQQPTTYAASTAPAYQAPPA------ASGPYQP---------------AST 981

Qy          64 SAARARANHNANTIRRTAMFAETD---PMTWLRPTVGLKRTFNPRIIRPQPPNP------ 114
               | | |         : |  : :::   | |  :|: |    : |   ||  | |
Db         982 STAYAPP-------QPTPSYGDSNPYAPATVYQPSSG----YPPNGYRPSDPQPQGYGAP 1030

Qy         115 ----------------SMSLGISG----PTILPQKTQSADQ-----SALQQP------- 141
                             |   ||||    ||  |:: ||| :     ||:  |
Db        1031 QPSFGQTQAMPPPPRVGSQQRGISGWNDAPTFAPKRPQSAAKDVKKASAILSPFPNSADP 1090

Qy         142 -----AALAFSGSSPQ----------HPPPQTT----------SASVGQQQHVVSGSSG 175
                  | |  :|::|           |||:                :| |||     :
Db        1091 LAAAGAGLNTAGAAPPPGRSPQPGVIPPPPKNARPPSIAAKVQPPPTVQQQQQFHQHAQH 1150

Qy         176 QQPQQGAQSSTVQPTTGSPPAAQGVPQST---PPP 207
               || |:  | :       ||||| | | |    |||
Db        1151 QQQQRQQQLT-------SPPAAAGPPPSAFSRPPP 1178
```

RESULT 4
US-10-437-963-166255
; Sequence 166255, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 166255
; LENGTH: 574
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_64982C.1.pep
US-10-437-963-166255

Query Match 10.9%; Score 131; DB 16; Length 574;
Best Local Similarity 28.1%; Pred. No. 0.22;
Matches 52; Conservative 15; Mismatches 64; Indels 54; Gaps 8;

```
Qy          88 PMTWLRPTVGLKRTFNPRIIRP---------QPPNPSMSLGISGPTILPQKTQSADQ--- 135
               | | | ||:  :  : |  ::|          || |          |  ||  |:  |
Db         307 PPTPLAPTIPQESYYPPSAVQPTDTTHQQYQAPPAPQS----QAPPAPPQHYQTPPQYAQ 362

Qy         136 ----------------SALQQPAALAFSGSSPQHPPPQTTSASVGQQQHVVSGSSGQQP 178
                             |  |||  :|    :|  ||||:   :|       |   ||  |
Db         363 YSQPPPASANPSTAVPPSVHQQPEEVA----APYGPPPQSYPPNVRLPSPYVPPPSGPAP 418

Qy         179 QQGAQSSTVQPTTGSPPAAQGVPQSTPPPTQNT---PQGGKG--------QTLSHTGQSG 227
                  :    |    ||| :  | | ||| :  ||   |||| |          : || | :|
Db         419 PFYGPN----PGMYEPPAVR--PNSGPPPSYNTGYKPQGGGGFPEPYGYSGSPSHRGNAG 472

Qy         228 NASRS 232
                 | |
```

```
Db        473 MKSPS 477


RESULT 5
US-10-425-114-58581
; Sequence 58581, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
;  APPLICANT: Liu, Jingdong
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Screen, Steven E
;  APPLICANT:  Tabaska, Jack E
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53313)B
;  CURRENT APPLICATION NUMBER: US/10/425,114
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 58581
;   LENGTH: 717
;   TYPE: PRT
;   ORGANISM: Zea mays
;   FEATURE:
;   OTHER INFORMATION: Clone ID: UC-ZMFLB73099H11_FLI.pep
US-10-425-114-58581

  Query Match             10.8%;  Score 130;  DB 15;  Length 717;
  Best Local Similarity   28.7%;  Pred. No. 0.34;
  Matches   50;  Conservative   15;  Mismatches   57;  Indels   52;  Gaps    7;

Qy         93 RPTVGLKRTFNPRIIRPQPPNPSMSLGISGPTILPQKTQSADQSALQQPAALAFSGSSPQ 152
              :| :| :: :      || |:|      |||   | :   |   : | |||
Db        489 QPPMGPQQGY------PQQPDPYARPPYSGPGQWPPRGAPAADGSYQAPAA--------S 534

Qy        153 HPPPQTTSASVGQQQHVVSGSSG--QQ--PQQGAQSSTVQPTTGSPPAAQGVPQ-----S 203
              : ||    :: ||     :|| |  ||  |||| |:    |   | ||| | ||     :
Db        535 YGPPSQHPSAYGQTYSAATGSDGYAQQGYPQQGGQAPA--PYGQSAPAAPGYPQQQGGYA 592

Qy        204 TPPPTQNT--------------------------PQGGKGQTLSHTGQSGNAS 230
                | ||                          || | |     |||:| ||
Db        593 QYPQTQPAYGEQAAQTNASYGYQGAPADPNYGSAYPQPGYGPPGQATGQAGYAS 646


RESULT 6
US-09-825-751A-79
; Sequence 79, Application US/09825751A
; Publication No. US20030065140A1
; GENERAL INFORMATION:
;  APPLICANT: CuraGen Corporation
;  APPLICANT:  Vernet, Corine A.M.
;  APPLICANT:  Fernandes, Elma R
;  APPLICANT:  Taupier, Raymond J
;  APPLICANT:  Quinn, Kerry E
;  APPLICANT:  Spytek, Kimberly A
;  APPLICANT:  Rastelli, Luca
;  APPLICANT:  Herrman, John L
;  TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
;  FILE REFERENCE: 15966-750
;  CURRENT APPLICATION NUMBER: US/09/825,751A
;  CURRENT FILING DATE:  2001-04-30
;  PRIOR APPLICATION NUMBER: 60/194,314
;  PRIOR FILING DATE: 2000-04-03
;  PRIOR APPLICATION NUMBER: 60/225,693
;  PRIOR FILING DATE: 2000-08-16
;  NUMBER OF SEQ ID NOS: 85
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 79
;   LENGTH: 1151
;   TYPE: PRT
;   ORGANISM: Gallus gallus
US-09-825-751A-79

  Query Match             10.6%;  Score 128;  DB 10;  Length 1151;
  Best Local Similarity   28.2%;  Pred. No. 0.8;
  Matches   57;  Conservative   22;  Mismatches   81;  Indels   42;  Gaps   11;

Qy         59 VGAA---------ASAARARANHNANTIRRTAMFAETDPMTWLRPTVGLKRTFNPRIIRP 109
              :|||         |:| :  |    ||:: : | | | | |   |:||   |       |
Db        776 MGAAATTQPSPMGAAATQVTATSAGNTMQVSPMGAATPPQT---PSVGAATT-------P 825

Qy        110 QPPNPSMSLGISGPTILPQKTQSADQSAL-QQPAALAFSGSSPQHPP-----PQTT--SA 161
              ||      :  :  |       | :   |:  ||  :| : :: | ||      ||:|   |
Db        826 QPSPMGAATTLMSPMGAATTPQPSPMGAVTTQPPPMAATNTT-QPPPMAASTPQSTPMGA 884

Qy        162 SVGQQQHVVSGSSGQQPQQGAQSSTVQPT-TGSP---------PAAQGVPQ---STPPPT 208
              :   |   :  :: | |  || :    ||  |||         | ||  ||    :|||
Db        885 ATTTQSPPMGATTTQSPPMGASTPQAPPTVAGSPTPPPPIPPSPTAQTSPQPMSKSPPPD 944

Qy        209 QNTPQGGKGQTLSHTGQSGNAS 230
                  || |      |||
Db        945 PPKAPSAAAQT-SPAAHVANAS 965


RESULT 7
US-10-851-438-79
; Sequence 79, Application US/10851438
; Publication No. US20050153305A1
; GENERAL INFORMATION:
;  APPLICANT: CuraGen Corporation
;  APPLICANT:  Vernet, Corine A.M.
;  APPLICANT:  Fernandes, Elma R
;  APPLICANT:  Taupier, Raymond J
;  APPLICANT:  Quinn, Kerry E
;  APPLICANT:  Spytek, Kimberly A
;  APPLICANT:  Rastelli, Luca
;  APPLICANT:  Herrman, John L
;  TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
;  FILE REFERENCE: 15966-750
;  CURRENT APPLICATION NUMBER: US/10/851,438
;  CURRENT FILING DATE:  2004-05-21
;  PRIOR APPLICATION NUMBER: US/09/825,751
;  PRIOR FILING DATE: 2001-04-30
;  PRIOR APPLICATION NUMBER: 60/194,314
;  PRIOR FILING DATE: 2000-04-03
;  PRIOR APPLICATION NUMBER: 60/225,693
;  PRIOR FILING DATE: 2000-08-16
;  NUMBER OF SEQ ID NOS: 85
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 79
;   LENGTH: 1151
;   TYPE: PRT
;   ORGANISM: Gallus gallus
US-10-851-438-79

  Query Match             10.6%;  Score 128;  DB 18;  Length 1151;
  Best Local Similarity   28.2%;  Pred. No. 0.8;
  Matches   57;  Conservative   22;  Mismatches   81;  Indels   42;  Gaps   11;

Qy         59 VGAA---------ASAARARANHNANTIRRTAMFAETDPMTWLRPTVGLKRTFNPRIIRP 109
              :|||         |:| :  |    ||:: : | | | | |   |:||   |       |
Db        776 MGAAATTQPSPMGAAATQVTATSAGNTMQVSPMGAATPPQT---PSVGAATT-------P 825

Qy        110 QPPNPSMSLGISGPTILPQKTQSADQSAL-QQPAALAFSGSSPQHPP-----PQTT--SA 161
              ||      :  :  |       | :   |:  ||  :| : :: | ||      ||:|   |
Db        826 QPSPMGAATTLMSPMGAATTPQPSPMGAVTTQPPPMAATNTT-QPPPMAASTPQSTPMGA 884

Qy        162 SVGQQQHVVSGSSGQQPQQGAQSSTVQPT-TGSP---------PAAQGVPQ---STPPPT 208
              :   |   :  :: | |  || :    ||  |||         | ||  ||    :|||
Db        885 ATTTQSPPMGATTTQSPPMGASTPQAPPTVAGSPTPPPPIPPSPTAQTSPQPMSKSPPPD 944

Qy        209 QNTPQGGKGQTLSHTGQSGNAS 230
                  || |      |||
Db        945 PPKAPSAAAQT-SPAAHVANAS 965
```

```
RESULT 8
US-10-156-761-10907
; Sequence 10907, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
;  APPLICANT: OMURA, SATOSHI
;  APPLICANT:  IKEDA, HARUO
;  APPLICANT:  ISHIKAWA, JUN
;  APPLICANT:  HORIKAWA, HIROSHI
;  APPLICANT:  SHIBA, TADAYOSHI
;  APPLICANT:  SAKAKI, YOSHIYUKI
;  APPLICANT:  HATTORI, MASAHIRA
;  TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
;  FILE REFERENCE: 249-262
;  CURRENT APPLICATION NUMBER: US/10/156,761
;  CURRENT FILING DATE:  2002-05-29
;  PRIOR APPLICATION NUMBER: JP 2001-204089
;  PRIOR FILING DATE: 2001-05-30
;  PRIOR APPLICATION NUMBER: JP 2001-272697
;  PRIOR FILING DATE: 2001-08-02
;  NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 10907
;   LENGTH: 797
;   TYPE: PRT
;   ORGANISM: Streptomyces avermitilis
US-10-156-761-10907

  Query Match             10.6%;  Score 127.5;  DB 14;  Length 797;
  Best Local Similarity   29.3%;  Pred. No. 0.58;
  Matches   61;  Conservative   14;  Mismatches  104;  Indels   29;  Gaps    8;

Qy         36 LNASGPLQPGHRVDIADARSIYTVGAAASAARARA--NHNANTIRRTAMFAETDPMTWLR 93
              | |  ||     :  |  |      |||:||| ||  :   :  | |:     |  | |
Db        460 LRALRPLPSPSLLPPAATRPASPRPAAAAAARPRALPSPTTSAARATSPSMTLPPPVWRR 519

Qy         94 PTVGLKRTFNPRIIRPQPPNPSMSLGISGPTILPQKTQSADQSALQQPAALAFSGSSPQH 153
              |          |  || || ||     |  |  |  |: :  |:   ||      :||
Db        520 PA---------RSSRP-PPRPS-----SRRTASPTPTRCSVWSSSPSPATTRSVSASPTP 564

Qy        154 PPPQTTSASVGQQQHVVSGSSGQQPQQGAQSSTVQPTTGSPPAAQGVPQSTPP-PTQNTP 212
              || :    |    :  | ||     :  | |  |   || :   | | || |   |
Db        565 PPSRKPPRST---RRPPSASSTTSRSRRTTSRTSSSTRRRPPTS---PASPPPRPPSRTS 618

Qy        213 QGGKGQTLSHTGQSGNASR-----SRRV 235
                 | :  |  |||      |||
Db        619 SASSAAWTSRSTSSRRASRRASRTPRRV 646


RESULT 9
US-09-328-599A-1
; Sequence 1, Application US/09328599A
; Publication No. US20020119529A1
;  GENERAL INFORMATION:
;    APPLICANT:  MOND, James J. and Lees, Andrew
;    TITLE OF INVENTION:  Enhancement of B Cell Activation by
;    TITLE OF INVENTION:  Co-Ligation of Receptors for Antigen and Complement C3d
;    TITLE OF INVENTION:  Using EBV gp350/220 or EBV gp350/220 Peptide Adjuvants
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Finnegan, Henderson, Farabow, Garrett &
;      ADDRESSEE:  Dunner, L.L.P.
;      STREET:  1300 I Street, N.W., Suite 700
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/328,599A
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fordis, Jean B.
;      REGISTRATION NUMBER:  32,984
;      REFERENCE/DOCKET NUMBER:  04995.6025-00000
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)408-4000
;      TELEFAX:  (202)408-4400
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  907 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-328-599A-1

  Query Match             10.6%;  Score 127.5;  DB 9;  Length 907;
  Best Local Similarity   24.9%;  Pred. No. 0.67;
  Matches   64;  Conservative   26;  Mismatches   88;  Indels   79;  Gaps   14;

Qy          2 TQPASSRVVFDP--SNPT---TFSVEAIAAYTPVALIRLLNASGPLQPGHRVDIADARSI 56
              | | |:     |  ::||   |     |:   ||       ||:||
Db        577 TSPTSAVTTPTPNATSPTLGKTSPTSAVTTPTP-------NATGP--------------- 614

Qy         57 YTVGAAASAARARANHNANTIRRTAMFAETDPMTWLRPTVGLKRTFNPRIIRPQPPNPSM 116
              |||  :  | |  ||   |:  |:            ||          ::  || | :
Db        615 -TVGETSPQANA-TNH---TLGGTS-----------PT---------PVVTSQPKNATS 648

Qy        117 SLGISGPTILPQKTQSADQSALQQPAAL--AFSGSSPQHPP-------------PQTTSA 161
              ::      |    | |       | |  : | :|  | |               | | |
Db        649 AVTTGQHNITSSSTSSMSLRPSSNPETLSPSTSDNSTSHMPLLTSAHPTGGENITQVTPA 708

Qy        162 SVGQQQHVVSGSSGQQPQQGAQSSTVQP---TTGSPPAAQGVPQSTPPPTQNTPQGGKGQ 218
              |:    | || ||  :|: |  |     |   :| : |    | : |||    :||   ||
Db        709 SI--STHHVSTSS-PEPRPGTTSQASGPGNSSTSTKPGEVNVTKGTPPQNATSPQAPSGQ 765

Qy        219 -----TLSHTGQSGNAS 230
                  |:: ||   |::
Db        766 KTAVPTVTSTGGKANST 782


RESULT 10
US-10-322-281-428
; Sequence 428, Application US/10322281
; Publication No. US20040126762A1
; GENERAL INFORMATION:
;  APPLICANT: David W. Morris
;  APPLICANT:  Marc S. Malandro
;  TITLE OF INVENTION: Novel Compositions and Methods in Cancer
;  FILE REFERENCE: 529452001000
;  CURRENT APPLICATION NUMBER: US/10/322,281
;  CURRENT FILING DATE:  2002-12-17
;  NUMBER OF SEQ ID NOS: 866
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 428
;   LENGTH: 630
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-322-281-428

  Query Match             10.5%;  Score 127;  DB 16;  Length 630;
  Best Local Similarity   28.0%;  Pred. No. 0.48;
  Matches   73;  Conservative   18;  Mismatches   96;  Indels   74;  Gaps   15;

Qy          1 MTQPASSRVVFDPSNPTTFSVEAIAAYTPVALIRLLNASGPLQP------GHRVDIADAR 54
              |:||        |: | :|||       |     :  :| || |      | :  :| |:
Db        200 MSQPT-------PAQPQSFSVGQPQPPPPPVGGAVAQSSAPLPPFPGAATGPQPMMAAAQ 252

Qy         55 SIYTVGAAASAARARANHNANTIRRTAMFAETDPMTWLRPTVGLKRTFNP-RIIRPQPPN 113
                    ||          :   |: : ||     |     | ||      | :   |||
```

```
Db        253 PSQPQGAGPGGQTLPPTN--VTLAQPAMSLPPQP----GPAVGAPAAQQPQQFAYPQPQI 306

Qy        114 PSMSLGISGPTILPQKTQSADQSA-LQQPAALAFSGSSPQHP-PPQTTSAS--------V 163

Db        307 PPGHL-------LP--VQPSGQSEYLQQHVA----GLQPPSPAQPSSTGAADATAATLPV 353

Qy        164 GQQQHVVS------GSSGQQ---------PQQGAQSSTVQPTTGSP------------PA 196

Db        354 GTGQNASSVGAQLMGASSQPSEAMAPRTGPAQGGQVAPCQPTGVPPATVGGVVQPCLGPA 413

Qy        197 AQGVPQSTPPPTQNTPQGGKG 217

Db        414 GAGQPQSVPPP----QMGGSG 430


RESULT 11
US-10-322-281-432
; Sequence 432, Application US/10322281
; Publication No. US20040126762A1
; GENERAL INFORMATION:
;  APPLICANT: David W. Morris
;  APPLICANT:  Marc S. Malandro
;  TITLE OF INVENTION: Novel Compositions and Methods in Cancer
;  FILE REFERENCE: 529452001000
;  CURRENT APPLICATION NUMBER: US/10/322,281
;  CURRENT FILING DATE:  2002-12-17
;  NUMBER OF SEQ ID NOS: 866
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 432
;   LENGTH: 720
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-322-281-432

  Query Match             10.5%;  Score 127;  DB 16;  Length 720;
  Best Local Similarity   28.0%;  Pred. No. 0.56;
  Matches   73;  Conservative   18;  Mismatches   96;  Indels   74;  Gaps   15;

Qy          1 MTQPASSRVVFDPSNPTTFSVEAIAAYTPVALIRLLNASGPLQP------GHRVDIADAR 54

Db        200 MSQPT-------PAQPQSFSVGQPQPPPPPVGGAVAQSSAPLPPFPGAATGPQPMMAAAQ 252

Qy         55 SIYTVGAAASAARARANHNANTIRRTAMFAETDPMTWLRPTVGLKRTFNP-RIIRPQPPN 113

Db        253 PSQPQGAGPGGQTLPPTN--VTLAQPAMSLPPQP----GPAVGAPAAQQPQQFAYPQPQI 306

Qy        114 PSMSLGISGPTILPQKTQSADQSA-LQQPAALAFSGSSPQHP-PPQTTSAS--------V 163

Db        307 PPGHL-------LP--VQPSGQSEYLQQHVA----GLQPPSPAQPSSTGAADATAATLPV 353

Qy        164 GQQQHVVS------GSSGQQ---------PQQGAQSSTVQPTTGSP------------PA 196

Db        354 GTGQNASSVGAQLMGASSQPSEAMAPRTGPAQGGQVAPCQPTGVPPATVGGVVQPCLGPA 413

Qy        197 AQGVPQSTPPPTQNTPQGGKG 217

Db        414 GAGQPQSVPPP----QMGGSG 430


RESULT 12
US-10-245-752-54
; Sequence 54, Application US/10245752
; Publication No. US20030064473A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin
;  APPLICANT:  Eaton,Dan
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Grimaldi,J. Christopher
;  APPLICANT:  Gurney,Austin
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stephan,Jean-Phillippe
;  APPLICANT:  Watanbe,Colin
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang,Zemin
;  APPLICANT:  Fong,Sherman
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3630R1C66
;  CURRENT APPLICATION NUMBER: US/10/245,752
;  CURRENT FILING DATE:  2002-09-16
;  PRIOR APPLICATION NUMBER: 10/197942
;  PRIOR FILING DATE: 2002-07-18
;  PRIOR APPLICATION NUMBER: 60/059114
;  PRIOR FILING DATE: 1997-09-17
;  PRIOR APPLICATION NUMBER: 60/063046
;  PRIOR FILING DATE: 1997-10-24
;  PRIOR APPLICATION NUMBER: 60/065027
;  PRIOR FILING DATE: 1997-11-10
;  PRIOR APPLICATION NUMBER: 60/079689
;  PRIOR FILING DATE: 1998-03-27
;  PRIOR APPLICATION NUMBER: 60/086478
;  PRIOR FILING DATE: 1998-05-22
;  PRIOR APPLICATION NUMBER: 60/087607
;  PRIOR FILING DATE: 1998-06-02
;  PRIOR APPLICATION NUMBER: 60/089801
;  PRIOR FILING DATE: 1998-06-18
;  PRIOR APPLICATION NUMBER: 60/090557
;  PRIOR FILING DATE: 1998-06-24
;  PRIOR APPLICATION NUMBER: 60/090689
;  PRIOR FILING DATE: 1998-06-25
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 54
;   LENGTH: 839
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-10-245-752-54

  Query Match             10.5%;  Score 126.5;  DB 14;  Length 839;
  Best Local Similarity   25.5%;  Pred. No. 0.72;
  Matches   76;  Conservative   24;  Mismatches   97;  Indels  101;  Gaps   16;

Qy         17 TTFSVEAIAA-----YTPVALIR-----------LLNASGPLQP----GHRVDIADARS 55

Db        338 TQVTVEAVAAAGSPPRFPQSLYRGTVARGAGAGVVVKDAAAPSQPLRIQAQDPEFSDLNS 397

Qy         56 IYT--------------VGAAASAARARA------NHN------ANTIRRTAMFAETDP 88

Db        398 AITYRITNHSHFRMEGEVVLTTTTLAQAGAFYAEVEAHNTVTSGTATTVIEIQVSEQEPP 457

Qy         89 MTWLRPTVGLKRTFNPRIIRPQPPNPSM------SLGISGP-----TILPQKTQS----- 132

Db        458 STEAGGTTGPWTSTTSEV--PRPPEPSQGPSTTSSGGGTGPHPPSGTTLRPPTSSTPGGP 515

Qy        133 --ADQSALQQPAALAFSGSSPQHPPPQTT-------SASVGQQQHVVSGSSGQQPQQGAQ 183

Db        516 PGAENSTSHQPATP--GGDTAQTPKPGTSQPMPPGVGTSTSHQPATPSGGTAQTPEPGTS 573

Qy        184 ---------SSTVQPTTGSPPAAQ----GVPQSTPP---------PTQNTPQGGKGQT 219

Db        574 QPMPPSMGTSTSHQPATPGGGTAQTPEAGTSQPMPPGMGTSTSHQPT--TPGGGTAQT 629


RESULT 13
US-10-245-859-54
; Sequence 54, Application US/10245859
; Publication No. US20030064474A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin
;  APPLICANT:  Eaton,Dan
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Grimaldi,J. Christopher
;  APPLICANT:  Gurney,Austin
;  APPLICANT:  Smith,Victoria
```

```
;  APPLICANT:  Stephan,Jean-Phillippe
;  APPLICANT:  Watanbe,Colin
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang,Zemin
;  APPLICANT:  Fong,Sherman
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3630R1C78
;  CURRENT APPLICATION NUMBER: US/10/245,859
;  CURRENT FILING DATE:  2002-09-16
;  PRIOR APPLICATION NUMBER: 10/197942
;  PRIOR FILING DATE: 2002-07-18
;  PRIOR APPLICATION NUMBER: 60/059114
;  PRIOR FILING DATE: 1997-09-17
;  PRIOR APPLICATION NUMBER: 60/063046
;  PRIOR FILING DATE: 1997-10-24
;  PRIOR APPLICATION NUMBER: 60/065027
;  PRIOR FILING DATE: 1997-11-10
;  PRIOR APPLICATION NUMBER: 60/079689
;  PRIOR FILING DATE: 1998-03-27
;  PRIOR APPLICATION NUMBER: 60/086478
;  PRIOR FILING DATE: 1998-05-22
;  PRIOR APPLICATION NUMBER: 60/087607
;  PRIOR FILING DATE: 1998-06-02
;  PRIOR APPLICATION NUMBER: 60/089801
;  PRIOR FILING DATE: 1998-06-18
;  PRIOR APPLICATION NUMBER: 60/090557
;  PRIOR FILING DATE: 1998-06-24
;  PRIOR APPLICATION NUMBER: 60/090689
;  PRIOR FILING DATE: 1998-06-25
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 54
;   LENGTH: 839
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-10-245-859-54

  Query Match             10.5%;  Score 126.5;  DB 14;  Length 839;
  Best Local Similarity   25.5%;  Pred. No. 0.72;
  Matches   76;  Conservative   24;  Mismatches   97;  Indels  101;  Gaps   16;

Qy         17 TTFSVEAIAA-----YTPVALIR-----------LLNASGPLQP----GHRVDIADARS 55
              |  :|||:||       | :| |            : :|: | ||        : :|  |
Db        338 TQVTVEAVAAAGSPPRFPQSLYRGTVARGAGAGVVVKDAAAPSQPLRIQAQDPEFSDLNS 397

Qy         56 IYT---------------VGAAASAARARA------NHN------ANTIRRTAMFAETDP 88
                |               |    : |:| |       ||      | |:    :  :  |
Db        398 AITYRITNHSHFRMEGEVVLTTTTLAQAGAFYAEVEAHNTVTSGTATTVIEIQVSEQEPP 457

Qy         89 MTWLRPTVGLKRTFNPRIIRPQPPNPSM------SLGISGP-----TILPQKTQS----- 132
               |    | |   :    :  |:|| ||       | | :||     | |   | |
Db        458 STEAGGTTGPWTSTTSEV--PRPPEPSQGPSTTSSGGGTGPHPPSGTTLRPPTSSTPGGP 515

Qy        133 --ADQSALQQPAALAFSGSSPQHPPPQTT-------SASVGQQQHVVSGSSGQQPQQGAQ 183
                |: |   |||     | : | | | |:          |   |    || : | |: |
Db        516 PGAENSTSHQPATP--GGDTAQTPKPGTSQPMPPGVGTSTSHQPATPSGGTAQTPEPGTS 573

Qy        184 ---------SSTVQPTTGSPPAAQ----GVPQSTPP---------PTQNTPQGGKGQT 219
                      |:: || |     ||    |  |  ||         ||  || ||  ||
Db        574 QPMPPSMGTSTSHQPATPGGGTAQTPEAGTSQPMPPGMGTSTSHQPT--TPGGGTAQT 629


RESULT 14
US-10-245-103-54
; Sequence 54, Application US/10245103
; Publication No. US20030068778A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin
;  APPLICANT:  Eaton,Dan
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Grimaldi,J. Christopher
;  APPLICANT:  Gurney,Austin
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stephan,Jean-Phillippe
;  APPLICANT:  Watanbe,Colin
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang,Zemin
;  APPLICANT:  Fong,Sherman
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3630R1C112
;  CURRENT APPLICATION NUMBER: US/10/245,103
;  CURRENT FILING DATE:  2002-09-17
;  PRIOR APPLICATION NUMBER: 10/197942
;  PRIOR FILING DATE: 2002-07-18
;  PRIOR APPLICATION NUMBER: 60/059114
;  PRIOR FILING DATE: 1997-09-17
;  PRIOR APPLICATION NUMBER: 60/063046
;  PRIOR FILING DATE: 1997-10-24
;  PRIOR APPLICATION NUMBER: 60/065027
;  PRIOR FILING DATE: 1997-11-10
;  PRIOR APPLICATION NUMBER: 60/079689
;  PRIOR FILING DATE: 1998-03-27
;  PRIOR APPLICATION NUMBER: 60/086478
;  PRIOR FILING DATE: 1998-05-22
;  PRIOR APPLICATION NUMBER: 60/087607
;  PRIOR FILING DATE: 1998-06-02
;  PRIOR APPLICATION NUMBER: 60/089801
;  PRIOR FILING DATE: 1998-06-18
;  PRIOR APPLICATION NUMBER: 60/090557
;  PRIOR FILING DATE: 1998-06-24
;  PRIOR APPLICATION NUMBER: 60/090689
;  PRIOR FILING DATE: 1998-06-25
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 54
;   LENGTH: 839
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-10-245-103-54

  Query Match             10.5%;  Score 126.5;  DB 14;  Length 839;
  Best Local Similarity   25.5%;  Pred. No. 0.72;
  Matches   76;  Conservative   24;  Mismatches   97;  Indels  101;  Gaps   16;

Qy         17 TTFSVEAIAA-----YTPVALIR-----------LLNASGPLQP----GHRVDIADARS 55
              |  :|||:||       | :| |            : :|: | ||        : :|  |
Db        338 TQVTVEAVAAAGSPPRFPQSLYRGTVARGAGAGVVVKDAAAPSQPLRIQAQDPEFSDLNS 397

Qy         56 IYT---------------VGAAASAARARA------NHN------ANTIRRTAMFAETDP 88
                |               |    : |:| |       ||      | |:    :  :  |
Db        398 AITYRITNHSHFRMEGEVVLTTTTLAQAGAFYAEVEAHNTVTSGTATTVIEIQVSEQEPP 457

Qy         89 MTWLRPTVGLKRTFNPRIIRPQPPNPSM------SLGISGP-----TILPQKTQS----- 132
               |    | |   :    :  |:|| ||       | | :||     | |   | |
Db        458 STEAGGTTGPWTSTTSEV--PRPPEPSQGPSTTSSGGGTGPHPPSGTTLRPPTSSTPGGP 515

Qy        133 --ADQSALQQPAALAFSGSSPQHPPPQTT-------SASVGQQQHVVSGSSGQQPQQGAQ 183
                |: |   |||     | : | | | |:          |   |    || : | |: |
Db        516 PGAENSTSHQPATP--GGDTAQTPKPGTSQPMPPGVGTSTSHQPATPSGGTAQTPEPGTS 573

Qy        184 ---------SSTVQPTTGSPPAAQ----GVPQSTPP---------PTQNTPQGGKGQT 219
                      |:: || |     ||    |  |  ||         ||  || ||  ||
Db        574 QPMPPSMGTSTSHQPATPGGGTAQTPEAGTSQPMPPGMGTSTSHQPT--TPGGGTAQT 629


RESULT 15
US-10-245-107-54
; Sequence 54, Application US/10245107
; Publication No. US20030068779A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin
;  APPLICANT:  Eaton,Dan
;  APPLICANT:  Filvaroff,Ellen
```

```
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Grimaldi,J. Christopher
;  APPLICANT:  Gurney,Austin
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stephan,Jean-Phillippe
;  APPLICANT:  Watanbe,Colin
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang,Zemin
;  APPLICANT:  Fong,Sherman
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3630R1C71
;  CURRENT APPLICATION NUMBER: US/10/245,107
;  CURRENT FILING DATE:  2002-09-16
;  PRIOR APPLICATION NUMBER: 10/197942
;  PRIOR FILING DATE: 2002-07-18
;  PRIOR APPLICATION NUMBER: 60/059114
;  PRIOR FILING DATE: 1997-09-17
;  PRIOR APPLICATION NUMBER: 60/063046
;  PRIOR FILING DATE: 1997-10-24
;  PRIOR APPLICATION NUMBER: 60/065027
;  PRIOR FILING DATE: 1997-11-10
;  PRIOR APPLICATION NUMBER: 60/079689
;  PRIOR FILING DATE: 1998-03-27
;  PRIOR APPLICATION NUMBER: 60/086478
;  PRIOR FILING DATE: 1998-05-22
;  PRIOR APPLICATION NUMBER: 60/087607
;  PRIOR FILING DATE: 1998-06-02
;  PRIOR APPLICATION NUMBER: 60/089801
;  PRIOR FILING DATE: 1998-06-18
;  PRIOR APPLICATION NUMBER: 60/090557
;  PRIOR FILING DATE: 1998-06-24
;  PRIOR APPLICATION NUMBER: 60/090689
;  PRIOR FILING DATE: 1998-06-25
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 54
;   LENGTH: 839
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-10-245-107-54

  Query Match             10.5%;  Score 126.5;  DB 14;  Length 839;
  Best Local Similarity   25.5%;  Pred. No. 0.72;
  Matches   76;  Conservative   24;  Mismatches   97;  Indels  101;  Gaps   16;

Qy         17 TTFSVEAIAA-----YTPVALIR------------LLNASGPLQP----GHRVDIADARS 55
              |  :|||:||        | :| |             : :|: | ||        : :|  |
Db        338 TQVTVEAVAAAGSPPRFPQSLYRGTVARGAGAGVVVKDAAAPSQPLRIQAQDPEFSDLNS 397

Qy         56 IYT---------------VGAAASAARARA------NHN------ANTIRRTAMFAETDP 88
                |               |    : |:| |        ||       | |:    :  :  |
Db        398 AITYRITNHSHFRMEGEVVLTTTTLAQAGAFYAEVEAHNTVTSGTATTVIEIQVSEQEPP 457

Qy         89 MTWLRPTVGLKRTFNPRIIRPQPPNPSM------SLGISGP-----TILPQKTQS----- 132
              |    | |   :    :  |:|| ||        | | :||     | |   | |
Db        458 STEAGGTTGPWTSTTSEV--PRPPEPSQGPSTTSSGGGTGPHPPSGTTLRPPTSSTPGGP 515

Qy        133 --ADQSALQQPAALAFSGSSPQHPPPQTT-------SASVGQQQHVVSGSSGQQPQQGAQ 183
                |: |   |||      | : | | | |:          |   |    || : | |: |
Db        516 PGAENSTSHQPATP--GGDTAQTPKPGTSQPMPPGVGTSTSHQPATPSGGTAQTPEPGTS 573

Qy        184 ---------SSTVQPTTGSPPAAQ----GVPQSTPP---------PTQNTPQGGKGQT 219
                      |:: || |      ||    |  | ||          ||  || ||  ||
Db        574 QPMPPSMGTSTSHQPATPGGGTAQTPEAGTSQPMPPGMGTSTSHQPT--TPGGGTAQT 629


Search completed: September 25, 2005, 21:42:37
Job time : 509 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: September 25, 2005, 21:05:32 ; Search time 3605 Seconds
(without alignments)
9475.985 Million cell updates/sec

Title: US-09-874-140-1
Perfect score: 705
Sequence: 1 atgacacaacccgcatcgtc.....cttcaagaagtcgtagggtg 705

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	705	100.0	705	6	A94878	A94878 Sequence 2
	2	705	100.0	705	6	AR161681	AR161681 Sequence
	3	705	100.0	705	6	E36500	E36500 Varicella-z
c	4	705	100.0	124883	14	AY548170	AY548170 Human her
c	5	705	100.0	124884	6	AR283580	AR283580 Sequence
c	6	705	100.0	124884	6	AR431491	AR431491 Sequence
c	7	705	100.0	124884	6	BD097170	BD097170 Method fo
c	8	705	100.0	124884	14	HEVZVXX	X04370 Human herpe
c	9	705	100.0	125459	14	AY548171	AY548171 Human her
c	10	703.4	99.8	125078	14	AB097932	AB097932 Human her
c	11	703.4	99.8	125125	14	AB097933	AB097933 Human her
c	12	703.4	99.8	125157	6	AR431492	AR431492 Sequence
c	13	703.4	99.8	125157	6	BD097171	BD097171 Method fo
c	14	181.2	25.7	124138	14	AF275348	AF275348 Cercopith
	15	40.4	5.7	151102	14	AF282130	AF282130 Meleagrid
	16	40.4	5.7	159160	14	AF291866	AF291866 Meleagrid
	17	40	5.7	125020	9	AF429315	AF429315 Homo sapi
	18	39.6	5.6	142533	2	AC138160	AC138160 Rattus no
c	19	39.6	5.6	189442	10	AC119382	AC119382 Rattus no
c	20	39.2	5.6	161264	1	AY456696	AY456696 Arthrobac
	21	38.8	5.5	232869	10	AL603925	AL603925 Mouse DNA
	22	38.6	5.5	34971	9	HS367G8	Z97634 Human DNA s
c	23	38.6	5.5	133069	2	AC026836	AC026836 Homo sapi
	24	38.6	5.5	260418	9	AE006463	AE006463 Homo sapi
c	25	38.4	5.4	2000	6	AX655393	AX655393 Sequence
	26	38.4	5.4	2167	4	AY035849	AY035849 Sus scrof
c	27	38.2	5.4	125020	9	AF429315	AF429315 Homo sapi
	28	37.8	5.4	11166	2	AC014870	AC014870 Drosophil
c	29	37.8	5.4	145597	14	AF030027	AF030027 Equine he
	30	37.8	5.4	151500	3	AC008355	AC008355 Drosophil
	31	37.8	5.4	159479	3	AC009735	AC009735 Drosophil
	32	37.8	5.4	299749	3	AE003606	AE003606 Drosophil
	33	37.4	5.3	124139	8	AC134930	AC134930 Oryza sat
	34	37.2	5.3	3030	8	VCA429230	AJ429230 Volvox ca
c	35	37.2	5.3	87061	2	AC142038	AC142038 Rattus no
	36	37.2	5.3	87076	9	AC005918	AC005918 Homo sapi
	37	37.2	5.3	178288	9	AC147017	AC147017 Pan trogl
c	38	37	5.2	302614	1	AE016778	AE016778 Pseudomon
	39	36.8	5.2	300885	1	AE012559	AE012559 Xylella f
c	40	36.6	5.2	922	8	AK071477	AK071477 Oryza sat
	41	36.6	5.2	1882	8	AK106923	AK106923 Oryza sat
	42	36.6	5.2	66237	2	AC109207	AC109207 Mus muscu
c	43	36.6	5.2	185080	9	AL954258	AL954258 Pan trogl
c	44	36.4	5.2	110000	2	LMFLCHR36_17	Continuation (18 o
	45	36.2	5.1	110000	1	AE000516_27	Continuation (28 o

ALIGNMENTS

RESULT 1
A94878
LOCUS A94878 705 bp DNA linear PAT 26-JAN-2000
DEFINITION Sequence 2 from Patent EP0927759.
ACCESSION A94878
VERSION A94878.1 GI:6779075
KEYWORDS .
SOURCE Human herpesvirus 3
ORGANISM Human herpesvirus 3
Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
Alphaherpesvirinae; Varicellovirus.
REFERENCE 1 (bases 1 to 705)
AUTHORS Giesendorf,B.D. and Eickmann,M.D.
TITLE Immunoreactive protein VP26 of Varicella Zoster Virus (VZV) and its diagnostic use
JOURNAL Patent: EP 0927759-A 2 07-JUL-1999;
DADE BEHRING MARBURG GMBH (DE)
FEATURES Location/Qualifiers
source 1. .705
/organism="Human herpesvirus 3"
/mol_type="unassigned DNA"
/strain="ELLEN"
/db_xref="taxon:10335"
/map="ORF23"
ORIGIN

Query Match 100.0%; Score 705; DB 6; Length 705;
Best Local Similarity 100.0%; Pred. No. 6.2e-205;
Matches 705; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ATGACACAACCCGCATCGTCTCGTGTAGTCTTTGATCCCAGCAACCCCACCACATTTTCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGACACAACCCGCATCGTCTCGTGTAGTCTTTGATCCCAGCAACCCCACCACATTTTCG 60

Qy         61 GTGGAAGCAATTGCGGCTTACACCCCCGTTGCTTTAATACGACTTTTAAACGCCAGTGGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTGGAAGCAATTGCGGCTTACACCCCCGTTGCTTTAATACGACTTTTAAACGCCAGTGGA 120

Qy        121 CCTTTGCAACCTGGTCACCGTGTGGACATCGCTGATGCCAGAAGCATTTACACCGTGGGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCTTTGCAACCTGGTCACCGTGTGGACATCGCTGATGCCAGAAGCATTTACACCGTGGGA 180
```

```
Qy        181 GCCGCGGCCAGTGCCGCGCGTGCACGCGCTAACCATAATGCAAATACGATACGCCGAACG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCCGCGGCCAGTGCCGCGCGTGCACGCGCTAACCATAATGCAAATACGATACGCCGAACG 240

Qy        241 GCCATGTTTGCCGAGACTGACCCTATGACATGGTTAAGACCAACGGTTGGCTTAAAACGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCCATGTTTGCCGAGACTGACCCTATGACATGGTTAAGACCAACGGTTGGCTTAAAACGT 300

Qy        301 ACGTTTAACCCGCGTATTATACGACCACAACCCCCAAATCCATCCATGAGTTTGGGAATC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ACGTTTAACCCGCGTATTATACGACCACAACCCCCAAATCCATCCATGAGTTTGGGAATC 360

Qy        361 TCGGGGCCTACTATATTGCCGCAAAAAACACAGAGCGCCGATCAGTCTGCTTTACAACAG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TCGGGGCCTACTATATTGCCGCAAAAAACACAGAGCGCCGATCAGTCTGCTTTACAACAG 420

Qy        421 CCCGCCGCGTTGGCGTTTTCGGGATCATCCCCGCAACACCCCCCACCTCAAACGACGTCG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CCCGCCGCGTTGGCGTTTTCGGGATCATCCCCGCAACACCCCCCACCTCAAACGACGTCG 480

Qy        481 GCATCCGTTGGACAACAGCAACACGTGGTGTCGGGGTCTTCTGGACAACAACCGCAACAG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCATCCGTTGGACAACAGCAACACGTGGTGTCGGGGTCTTCTGGACAACAACCGCAACAG 540

Qy        541 GGAGCACAGTCAAGCACTGTCCAGCCAACAACCGGATCACCGCCCGCGGCCCAAGGCGTG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GGAGCACAGTCAAGCACTGTCCAGCCAACAACCGGATCACCGCCCGCGGCCCAAGGCGTG 600

Qy        601 CCACAGTCTACCCCGCCCCCAACCCAAAATACCCCCCAGGGGGGTAAGGGACAGACCTTG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CCACAGTCTACCCCGCCCCCAACCCAAAATACCCCCCAGGGGGGTAAGGGACAGACCTTG 660

Qy        661 TCACACACGGGACAATCTGGAAACGCTTCAAGAAGTCGTAGGGTG 705
              |||||||||||||||||||||||||||||||||||||||||||||
Db        661 TCACACACGGGACAATCTGGAAACGCTTCAAGAAGTCGTAGGGTG 705
```

RESULT 2
AR161681

```
LOCUS       AR161681                705 bp    DNA     linear   PAT 17-OCT-2001
DEFINITION  Sequence 1 from patent US 6258363.
ACCESSION   AR161681
VERSION     AR161681.1  GI:16228540
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 705)
  AUTHORS   Eickmann,M., Gicklhorn,D., Radsak,K., Hauser,H.-P. and
            Giesendorf,B.
  TITLE     Varicella zoster virus (VZV) immunoreactive protein VP26 and its
            diagnostic use
  JOURNAL   Patent: US 6258363-A 1 10-JUL-2001;
FEATURES             Location/Qualifiers
     source          1. .705
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN
```

```
  Query Match           100.0%;  Score 705;  DB 6;  Length 705;
  Best Local Similarity  100.0%;  Pred. No. 6.2e-205;
  Matches  705;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGACACAACCCGCATCGTCTCGTGTAGTCTTTGATCCCAGCAACCCCACCACATTTTCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGACACAACCCGCATCGTCTCGTGTAGTCTTTGATCCCAGCAACCCCACCACATTTTCG 60

Qy         61 GTGGAAGCAATTGCGGCTTACACCCCCGTTGCTTTAATACGACTTTTAAACGCCAGTGGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTGGAAGCAATTGCGGCTTACACCCCCGTTGCTTTAATACGACTTTTAAACGCCAGTGGA 120

Qy        121 CCTTTGCAACCTGGTCACCGTGTGGACATCGCTGATGCCAGAAGCATTTACACCGTGGGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCTTTGCAACCTGGTCACCGTGTGGACATCGCTGATGCCAGAAGCATTTACACCGTGGGA 180

Qy        181 GCCGCGGCCAGTGCCGCGCGTGCACGCGCTAACCATAATGCAAATACGATACGCCGAACG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCCGCGGCCAGTGCCGCGCGTGCACGCGCTAACCATAATGCAAATACGATACGCCGAACG 240

Qy        241 GCCATGTTTGCCGAGACTGACCCTATGACATGGTTAAGACCAACGGTTGGCTTAAAACGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCCATGTTTGCCGAGACTGACCCTATGACATGGTTAAGACCAACGGTTGGCTTAAAACGT 300

Qy        301 ACGTTTAACCCGCGTATTATACGACCACAACCCCCAAATCCATCCATGAGTTTGGGAATC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ACGTTTAACCCGCGTATTATACGACCACAACCCCCAAATCCATCCATGAGTTTGGGAATC 360

Qy        361 TCGGGGCCTACTATATTGCCGCAAAAAACACAGAGCGCCGATCAGTCTGCTTTACAACAG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TCGGGGCCTACTATATTGCCGCAAAAAACACAGAGCGCCGATCAGTCTGCTTTACAACAG 420

Qy        421 CCCGCCGCGTTGGCGTTTTCGGGATCATCCCCGCAACACCCCCCACCTCAAACGACGTCG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CCCGCCGCGTTGGCGTTTTCGGGATCATCCCCGCAACACCCCCCACCTCAAACGACGTCG 480

Qy        481 GCATCCGTTGGACAACAGCAACACGTGGTGTCGGGGTCTTCTGGACAACAACCGCAACAG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCATCCGTTGGACAACAGCAACACGTGGTGTCGGGGTCTTCTGGACAACAACCGCAACAG 540

Qy        541 GGAGCACAGTCAAGCACTGTCCAGCCAACAACCGGATCACCGCCCGCGGCCCAAGGCGTG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GGAGCACAGTCAAGCACTGTCCAGCCAACAACCGGATCACCGCCCGCGGCCCAAGGCGTG 600

Qy        601 CCACAGTCTACCCCGCCCCCAACCCAAAATACCCCCCAGGGGGGTAAGGGACAGACCTTG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CCACAGTCTACCCCGCCCCCAACCCAAAATACCCCCCAGGGGGGTAAGGGACAGACCTTG 660

Qy        661 TCACACACGGGACAATCTGGAAACGCTTCAAGAAGTCGTAGGGTG 705
              |||||||||||||||||||||||||||||||||||||||||||||
Db        661 TCACACACGGGACAATCTGGAAACGCTTCAAGAAGTCGTAGGGTG 705
```

RESULT 3
E36500

```
LOCUS       E36500                  705 bp    DNA     linear   PAT 18-JUN-2001
DEFINITION  Varicella-zoster virus (VZV) immune reaction protein VP26 and
            diagnostic utilization thereof.
ACCESSION   E36500
VERSION     E36500.1  GI:13022697
KEYWORDS    JP 1999243974-A/1.
SOURCE      Human herpesvirus 3
  ORGANISM  Human herpesvirus 3
            Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
            Alphaherpesvirinae; Varicellovirus.
REFERENCE   1  (bases 1 to 705)
  AUTHORS   Markus,A., Dorotia,G., Klaus,R., Hanspater,H. and Bernhard,G.
  TITLE     Varicella-zoster virus (VZV) immune reaction protein VP26 and
            diagnostic utilization thereof
  JOURNAL   Patent: JP 1999243974-A 1 14-SEP-1999;
            DEIDO BERINGU MARUBURUKU GMBH
COMMENT     OS   Varicella-zoster virus
            PN   JP 1999243974-A/1
            PD   14-SEP-1999
            PF   22-DEC-1998 JP 1998364934
            PR   23-DEC-1997 DE   19757765:2
            PI   MARKUS AIKUMAN,DOROTIA GIKURUHORUN,KLAUS RATOZAKU, PI
            HANS-PATER HAUSER,
            PI   BERNHARD  GIZENDORUFU
            PC   C12N15/09,C07K14/04,C07K16/08,C12Q1/68,G01N33/53,G01N33/569//
            PC   (C12N15/09,C12R1:92),C12N15/00,(C12N15/00,C12R1:92) CC
            Strandedness: Single;
            CC   Topology: Linear;
            FH   Key             Location/Qualifiers
            FT   source          1. .705
            FT                   /organism='Varicella-zoster virus'.
FEATURES             Location/Qualifiers
```

```
     source          1. .705
                     /organism="Human herpesvirus 3"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10335"
ORIGIN

  Query Match           100.0%;  Score 705;  DB 6;  Length 705;
  Best Local Similarity   100.0%;  Pred. No. 6.2e-205;
  Matches  705;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGACACAACCCGCATCGTCTCGTGTAGTCTTTGATCCCAGCAACCCCACCACATTTTCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGACACAACCCGCATCGTCTCGTGTAGTCTTTGATCCCAGCAACCCCACCACATTTTCG 60

Qy         61 GTGGAAGCAATTGCGGCTTACACCCCCGTTGCTTTAATACGACTTTTAAACGCCAGTGGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTGGAAGCAATTGCGGCTTACACCCCCGTTGCTTTAATACGACTTTTAAACGCCAGTGGA 120

Qy        121 CCTTTGCAACCTGGTCACCGTGTGGACATCGCTGATGCCAGAAGCATTTACACCGTGGGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCTTTGCAACCTGGTCACCGTGTGGACATCGCTGATGCCAGAAGCATTTACACCGTGGGA 180

Qy        181 GCCGCGGCCAGTGCCGCGCGTGCACGCGCTAACCATAATGCAAATACGATACGCCGAACG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCCGCGGCCAGTGCCGCGCGTGCACGCGCTAACCATAATGCAAATACGATACGCCGAACG 240

Qy        241 GCCATGTTTGCCGAGACTGACCCTATGACATGGTTAAGACCAACGGTTGGCTTAAAACGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCCATGTTTGCCGAGACTGACCCTATGACATGGTTAAGACCAACGGTTGGCTTAAAACGT 300

Qy        301 ACGTTTAACCCGCGTATTATACGACCACAACCCCCAAATCCATCCATGAGTTTGGGAATC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ACGTTTAACCCGCGTATTATACGACCACAACCCCCAAATCCATCCATGAGTTTGGGAATC 360

Qy        361 TCGGGGCCTACTATATTGCCGCAAAAAACACAGAGCGCCGATCAGTCTGCTTTACAACAG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TCGGGGCCTACTATATTGCCGCAAAAAACACAGAGCGCCGATCAGTCTGCTTTACAACAG 420

Qy        421 CCCGCCGCGTTGGCGTTTTCGGGATCATCCCCGCAACACCCCCCACCTCAAACGACGTCG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CCCGCCGCGTTGGCGTTTTCGGGATCATCCCCGCAACACCCCCCACCTCAAACGACGTCG 480

Qy        481 GCATCCGTTGGACAACAGCAACACGTGGTGTCGGGGTCTTCTGGACAACAACCGCAACAG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCATCCGTTGGACAACAGCAACACGTGGTGTCGGGGTCTTCTGGACAACAACCGCAACAG 540

Qy        541 GGAGCACAGTCAAGCACTGTCCAGCCAACAACCGGATCACCGCCCGCGGCCCAAGGCGTG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GGAGCACAGTCAAGCACTGTCCAGCCAACAACCGGATCACCGCCCGCGGCCCAAGGCGTG 600

Qy        601 CCACAGTCTACCCCGCCCCCAACCCAAAATACCCCCCAGGGGGGTAAGGGACAGACCTTG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CCACAGTCTACCCCGCCCCCAACCCAAAATACCCCCCAGGGGGGTAAGGGACAGACCTTG 660

Qy        661 TCACACACGGGACAATCTGGAAACGCTTCAAGAAGTCGTAGGGTG 705
              |||||||||||||||||||||||||||||||||||||||||||||
Db        661 TCACACACGGGACAATCTGGAAACGCTTCAAGAAGTCGTAGGGTG 705


RESULT 4
AY548170/c
LOCUS       AY548170              124883 bp    DNA     linear   VRL 14-JUN-2004
DEFINITION  Human herpesvirus 3 strain MSP, complete genome.
ACCESSION   AY548170 AH009871 AY005330 AY005331 AY005332 AY005333 AY005334
            AY005335 AY034034
VERSION     AY548170.1  GI:46981482
KEYWORDS    .
SOURCE      Human herpesvirus 3
  ORGANISM  Human herpesvirus 3
            Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
            Alphaherpesvirinae; Varicellovirus.
REFERENCE   1  (bases 1 to 124883)
  AUTHORS   Santos,R.A., Hatfield,C.C., Cole,N.L., Padilla,J.A., Moffat,J.F.,
            Arvin,A.M., Ruyechan,W.T., Hay,J. and Grose,C.
  TITLE     Varicella-zoster virus gE escape mutant VZV-MSP exhibits an
            accelerated cell-to-cell spread phenotype in both infected cell
            cultures and SCID-hu mice
  JOURNAL   Virology 275 (2), 306-317 (2000)
  MEDLINE   20455576
   PUBMED   10998331
REFERENCE   2  (bases 1 to 124883)
  AUTHORS   Grose,C., Tyler,S., Peters,G., Hiebert,J., Stephens,G.M.,
            Ruyechan,W.T., Jackson,W., Storlie,J. and Tipples,G.A.
  TITLE     Complete DNA Sequence Analyses of the First Two Varicella-Zoster
            Virus Glycoprotein E (D150N) Mutant Viruses Found in North America:
            Evolution of Genotypes with an Accelerated Cell Spread Phenotype
  JOURNAL   J. Virol. 78 (13), 6799-6807 (2004)
   PUBMED   15194755
REFERENCE   3  (bases 1 to 124883)
  AUTHORS   Santos,R.A., Hatfield,C.C., Cole,N.L., Padilla,J.A., Moffat,J.F.,
            Arvin,A.M., Ruyechan,W.T., Hay,J. and Grose,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-JUN-2000) Department of Microbiology, University of
            Iowa, Iowa City, IA 52242, USA
REFERENCE   4  (bases 1 to 124883)
  AUTHORS   Santos,R.A., Hatfield,C.C., Cole,N.L., Padilla,J.A., Moffat,J.F.,
            Arvin,A.M., Ruyechan,W.T., Hay,J. and Grose,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-MAY-2001) Department of Microbiology, University of
            Iowa, Iowa City, IA 52242, USA
  REMARK    Sequence update by submitter
REFERENCE   5  (bases 1 to 124883)
  AUTHORS   Grose,C., Tyler,S., Peters,G., Hiebert,J., Stephens,G.,
            Ruyechan,W., Jackson,W., Storlie,J. and Tipples,G.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-FEB-2004) Health Canada, National Microbiology
            Laboratory, 1015 Arlington Street, Winnipeg, Manitoba R3E 3P6,
            Canada
COMMENT     On or before May 4, 2004 this sequence version replaced
            gi:14030155, gi:10697346, gi:10697347, gi:13940791, gi:10697349,
            gi:10697350, gi:10697351, gi:14030846.
FEATURES             Location/Qualifiers
     source          1. .124883
                     /organism="Human herpesvirus 3"
                     /mol_type="genomic DNA"
                     /strain="MSP"
                     /db_xref="taxon:10335"
     CDS             complement(590. .916)
                     /note="hypothetical protein; HHV3g13"
                     /codon_start=1
                     /product="ORF1"
                     /protein_id="AAT07759.1"
                     /db_xref="GI:46981483"
                     /translation="MSRVSEYGVPEGVRESDSDTDSVFMYQHTELMQNNASPLVVQTR
                     PPAVLIPLVDVPRPRSRRKASAQLKMQMDRLCNVLGVVLQMATLALVTYIAFVVHTRA
                     TSCKRE"
     CDS             1135. .1851
                     /note="ORF2; HHV3g29"
                     /codon_start=1
                     /product="immediate-early protein"
                     /protein_id="AAT07760.1"
                     /db_xref="GI:46981484"
                     /translation="MHVISETLAYGHVPAFIMGSTLVRPSLNATAEENPASETRCLLR
                     VLAGRTVDLPGGGTLHITCTKTYVIIGKYSKPGERLSLARLIGRAMTPGGARTFIILA
                     MKEKRSTTLGYECGTGLHLLAPSMGTFLRTHGLSNRDLCLWRGNIYDMHMQRLMFWEN
                     IAQNTTETPCITSTLTCNLTEDSGEAALTTSDRPTLPTLTAQGRPTVSNIRGILKGSP
                     RQQPVCHRVRFAEPTEGVLM"
     CDS             complement(1909. .2448)
                     /note="ORF3; HHV3g30"
                     /codon_start=1
                     /evidence=not_experimental
                     /product="hypothetical protein"
                     /protein_id="AAT07761.1"
                     /db_xref="GI:46981485"
                     /translation="MDTTGASESSQPIRVNLKPDPLASFTQVIPPLALETTWTCPANS
                     HAPTPSPLYGVKRLCALRATCGRADDLHAFLIGLGRRDKPSESPMYVDLQPFCSLLNS
```

QRLLPEMANYNTLCDAPFSAATQQMMLESGQLGVHLAAIGYHCHCKSPFSAECWTGAS
EAYDHVVCGGKARAAVGGL"
CDS complement(2784. .4142)
/note="ORF4; HHV3g21"
/codon_start=1
/product="transactivator"
/protein_id="AAT07762.1"
/db_xref="GI:46981486"
/translation="MASASIPTDPDVSTICEDFMNLLPDEPSDDFALEVTDWANDEAI
GSTPGEDSTTSRTVYVERTADTAYNPRYSKRRHGRRESYHHNRPKTLVVVLPDSNHHG
GRDVETGYARIERGHRRSSRSYNTQSSRKHRDRSLSNRRRRPTTPPAMTTGERNDQTH
DESYRLRFSKRDARRERIRKEYDIPVDRITGRAIEVVSTAGASVTIDSVRHLDETIEK
LVVRYATIQEGDSWASGGCFPGIKQNTSWPELMLYGHELYRTFESYKMDSRIARALRE
RVIRGESLIEALESADELLTWIKMLAAKNLPIYTNNPIVATSKSLLENLKLKLGPFVR
CLLLNRDNDLGSRTLPELLRQQRFSDITCITTYMFVMIARIANIVVRGSKFVEYDDIS
CNVQVLQEYTPGSCLAGVLEALITHQRECGRVECTLSTWAGHLSDARPYGKYFKCSTF
NC"
CDS complement(4253. .5275)
/note="ORF5; HHV3g22"
/codon_start=1
/product="glycoprotein K"
/protein_id="AAT07763.1"
/db_xref="GI:46981487"
/translation="MQALGIKTEHFIIMCLLSGHAVFTLWYTARVKFEHECVYATTVI
NGGPVVWGSYNNSLIYVTFVNHSTFLDGLSGYDYSCRENLLSGDTMVKTAISTPLHDK
IRIVLGTRNCHAYFWCVQLKMIFFAWFVYGMYLQFRRIRRMFGPFRSSCELISPTSYS
LNYVTRVISNILLGYPYTKLARLLCDVSMRRDGMSKVFNADPISFLYMHKGVTLLMLL
EVIAHISSGCIVLLTLGVAYTPCALLYPTYIRILAWVVVCTLAIVELISYVRPKPTKD
NHLNHINTGGIRGICTTCCATVMSGLAIKCFYIVIFAIAVVIFMHYEQRVQVSLFGES
ENSQKH"
CDS complement(5327. .8578)
/note="ORF6; HHV3g25"
/codon_start=1
/product="DNA helicase-primase complex"
/protein_id="AAT07764.1"
/db_xref="GI:46981488"
/translation="MDKSSKPTIRLLFATKGCAISHSLLLLTGQISTEPLYVVSYTWT
PDLDDVFVKNGREEITQVIPTKRPREVTENDEENQIMHLFCSRDVNVIFYLIGGFSTG
DVRSRVWPIFFCCFKTQTDFKALYKALWYGAPLNPHIISDTLCISETFDIHSEVIQTL
MVTTHHLNRKGLSDNGLCITEATLCKLVKKSVGRQELTSLYAHYERQVLAAYRRLYWG
YGCSPFWYIVRFGPSEKTLVLATRYYLLQTDTSYNTLETPLYDLQAIKDLFLTYQVPA
LPNCSGYNISDLLSFDKLSMFCCSSTYTRGLTAKNALSYILQRIHTDTTEIHAVSEYI
TNDRKGLKVPDREFVDYIYLAHFECFNRKQIADHLQAVTYSDFVNKPVLLKSSNLGKR
ATANFFNHVRSRLNMRDYIKKNVICDVTELGPEIGHKYTITKTYTLSLTYAAKPSKFI
GVCDLATTLTRRVENIEKQFSPYGWSSTIPSNPPGFDELSNPEDSGVSAEALRAANFA
NDTPNQSGRTGFDTSPGITKLLLFFSAATGIATHDVSILSYKTPLEALIGHSEVTGPM
PVYRVALPHGAQAFAVIANDTWSSITNRYTLPHEARLIAEDLKQINPCNFVAASLRDM
QLTLLLSTSVKNVSKISSNIPKDQLYINRNELFNTNLIITNLILDVDFHIRKPIPLGI
LHAGMRAFRHGILTAMQLLFPKAVVNPNKDPCYFYKTACPEPTVEVLDDDNLLDITSH
SDIDFYIENGELYTCVEENYTEDVWFFDTQTTSEVHTHADVSNNENLHETLPCNCKEK
IGFRVCVPIPNPYALVGSSTLKGFAQILQQAVLLEREFVEYIGPYLRDFSFIDTGVYS
HGHSLRLPFFSKVTTTGTAVGQLLPFYVVPEQCIDILAFVTSHRNPANFHFHSRPQSN
VPVQFILHNLGGEYAEFFERKVARNKQIFSSPQISLTKALKERGVTCLDAFTLEAFVD
STILESIVEHIAVHFPGRDREYTLTSSKCIAIKRDWVLFQLICGTKGFTCLRYPHRGG
RTAPRTFVSLRVDHHNRLCISLAQQCFATKCDSNRMHTIFTLEVPNYPNLTSS"
CDS 8608. .9387
/note="ORF7; HHV3g26"
/codon_start=1
/product="virion phosphoprotein"
/protein_id="AAT07765.1"
/db_xref="GI:46981489"
/translation="MQTVCASLCGYARIPTEEPSYEEVRVNTHPQGAALLRLQEALTA
VNGLLPAPLTLEDVVASADNTRRLVRAQALARTYAACSRNIECLKQHHFTEDNPGLNA
VVRSHMENSKRLADMCLAAITHLYLSVGAVDVTTDDIVDQTLRMTAESEVVMSDVVLL
EKTLGVVAKPQASFDVSHNHELSIAKGENVGLKTSPIKSEATQLSEIKPPLIEVSDNN
TSNLTKKTYPTETLQPVLTPKQTQDVQRTTPAIKKSHVMLV"
CDS complement(9478. .10668)
/note="ORF8; HHV3g20"
/codon_start=1
/product="dUTPase"
/protein_id="AAT07766.1"
/db_xref="GI:46981490"
/translation="MNEAVIDPILETAVNTGDMFCSQTIPNRCLKDTILIEVQPECAD
TLQCVLDDKVSRHQPLLLRNHKKLELPSEKSVTRGGFYMQQLELLVKSAPPNEYALLL
IQCKDTALADEDNFFVANGVIDAGYRGVISALLYYRPGVTVILPGHLTIYLFPVKLRQ
SRLLPKNVLKHLDPIFKSIQVQPLSNSPSNYEKPVIPEFADISTVQQGQPLHRDSAEY
HIDVPLTYKHIINPKRQEDAGYDICVPYNLYLKRNEFIKIVLPIIRDWDLQHPSINAY
IFGRSSKSRSGIIVCPTAWPAGEHCKFYVYNLTGDDIRIKTGDRLAQVLLIDHNTQIH
LKHNVLSNIAFPYAIRGKCGIPGVQWYFTKTLDLIATPSERGTRGFGSTDKETNDVDF
LLKH"
CDS 11010. .11918
/note="ORF9; HHV3g23"
/codon_start=1
/product="tegument protein"
/protein_id="AAT07767.1"
/db_xref="GI:46981491"
/translation="MASSDGDRLCRSNAVRRKTTPSYSGQYRTARRSVVVGPPDDSDD
SLGYITTVGADSPSPVYADLYFEHKNTTPRVHQPNDSSGSEDDFEDIDEVVAAFREAR
LRHELVEDAVYENPLSVEKPSRSFTKNAAVKPKLEDSPKRAPPGAGAIASGRPISFST
APKTATSSWCGPTPSYNKRVFCEAVRRVAAMQAQKAAEAAWNSNPPRNNAELDRLLTG
AVIRITVHEGLNLIQAANEADLGEGASVSKRGHNRKTGDLQGGMGNEPMYAQVRKPKS
RTDTQTTGRITNRSRARSASRTDTRK"
CDS 12161. .13393
/note="ORF10; HHV3g24"
/codon_start=1
/product="transactivator"
/protein_id="AAT07768.1"
/db_xref="GI:46981492"
/translation="MECNLGTEHPSTDTWNRSKTEQAVVDAFDESLFGDVASDIGFET
SLYSHAVKTAPSPPWVASPKILYQQLIRDLDFSEGPRLLSCLETWNEDLFSCFPINED
LYSDMMVLSPDPDDVISTVSTKDHVEMFNLTTRGSVRLPSPPKQPTGLPAYVQEVQDS
FTVELRAREEAYTKLLVTYCKSIIRYLQGTAKRTTIGLNIQNPDQKAYTQLRQSILLR
YYREVASLARLLYLHLYLTVTREFSWRLYASQSAHPDVFAALKFTWTERRQFTCAFHP
VLCNHGIVLLEGKPLTASALREINYRRRELGLPLVRCGLVEENKSPLVQQPSFSVHLP
RSVGFLTHHIKRKLDAYAVKHPQEPRHVRADHPYAKVVENRNYGSSIEAMILAPPSPS
EILPGDPPRPPTCGFLTR"
CDS 13591. .16005
/note="ORF11; HHV3g03"
/codon_start=1
/product="tegument protein and transactivator"
/protein_id="AAT07769.1"
/db_xref="GI:46981493"
/translation="MQSGHYNRRQSRRQRISSNTTDSPRHTHGTRYRSTNWYTHPPQI
LSNSETLVAVQELLNSEMDQDSSSDASDDFPGYALHHSTYNGSEQNTSTSRHENRIFK
LTEREANEEININTDAIDDEGEAEEGEAEEDAIDDEGEAEEGEAEEDAIDDEGEAEEG

Query Match 100.0%; Score 705; DB 14; Length 124883;
Best Local Similarity 100.0%; Pred. No. 1.3e-204;
Matches 705; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGACACAACCCGCATCGTCTCGTGTAGTCTTTGATCCCAGCAACCCCACCACATTTTCG 60
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 43153 ATGACACAACCCGCATCGTCTCGTGTAGTCTTTGATCCCAGCAACCCCACCACATTTTCG 43094

Qy 61 GTGGAAGCAATTGCGGCTTACACCCCCGTTGCTTTAATACGACTTTTAAACGCCAGTGGA 120
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 43093 GTGGAAGCAATTGCGGCTTACACCCCCGTTGCTTTAATACGACTTTTAAACGCCAGTGGA 43034

Qy 121 CCTTTGCAACCTGGTCACCGTGTGGACATCGCTGATGCCAGAAGCATTTACACCGTGGGA 180
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 43033 CCTTTGCAACCTGGTCACCGTGTGGACATCGCTGATGCCAGAAGCATTTACACCGTGGGA 42974

Qy 181 GCCGCGGCCAGTGCCGCGCGTGCACGCGCTAACCATAATGCAAATACGATACGCCGAACG 240
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 42973 GCCGCGGCCAGTGCCGCGCGTGCACGCGCTAACCATAATGCAAATACGATACGCCGAACG 42914

Qy 241 GCCATGTTTGCCGAGACTGACCCTATGACATGGTTAAGACCAACGGTTGGCTTAAAACGT 300
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 42913 GCCATGTTTGCCGAGACTGACCCTATGACATGGTTAAGACCAACGGTTGGCTTAAAACGT 42854

Qy 301 ACGTTTAACCCGCGTATTATACGACCACAACCCCCAAATCCATCCATGAGTTTGGGAATC 360
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 42853 ACGTTTAACCCGCGTATTATACGACCACAACCCCCAAATCCATCCATGAGTTTGGGAATC 42794

Qy 361 TCGGGGCCTACTATATTGCCGCAAAAAACACAGAGCGCCGATCAGTCTGCTTTACAACAG 420
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

```
Db      42793 TCGGGGCCTACTATATTGCCGCAAAAAACACAGAGCGCCGATCAGTCTGCTTTACAACAG 42734

Qy        421 CCCGCCGCGTTGGCGTTTTCGGGATCATCCCCGCAACACCCCCCACCTCAAACGACGTCG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42733 CCCGCCGCGTTGGCGTTTTCGGGATCATCCCCGCAACACCCCCCACCTCAAACGACGTCG 42674

Qy        481 GCATCCGTTGGACAACAGCAACACGTGGTGTCGGGGTCTTCTGGACAACAACCGCAACAG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42673 GCATCCGTTGGACAACAGCAACACGTGGTGTCGGGGTCTTCTGGACAACAACCGCAACAG 42614

Qy        541 GGAGCACAGTCAAGCACTGTCCAGCCAACAACCGGATCACCGCCCGCGGCCCAAGGCGTG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42613 GGAGCACAGTCAAGCACTGTCCAGCCAACAACCGGATCACCGCCCGCGGCCCAAGGCGTG 42554

Qy        601 CCACAGTCTACCCCGCCCCCAACCCAAAATACCCCCCAGGGGGGTAAGGGACAGACCTTG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42553 CCACAGTCTACCCCGCCCCCAACCCAAAATACCCCCCAGGGGGGTAAGGGACAGACCTTG 42494

Qy        661 TCACACACGGGACAATCTGGAAACGCTTCAAGAAGTCGTAGGGTG 705
              |||||||||||||||||||||||||||||||||||||||||||||
Db      42493 TCACACACGGGACAATCTGGAAACGCTTCAAGAAGTCGTAGGGTG 42449
```

RESULT 5
AR283580/c
LOCUS AR283580 124884 bp DNA linear PAT 10-APR-2003
DEFINITION Sequence 76 from patent US 6528066.
ACCESSION AR283580
VERSION AR283580.1 GI:29720465
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 124884)
AUTHORS Grose,C.F. and Santos,R.
TITLE Variant varicella-zoster viruses and methods of use
JOURNAL Patent: US 6528066-A 76 04-MAR-2003;
FEATURES Location/Qualifiers
source 1. .124884
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN

Query Match 100.0%; Score 705; DB 6; Length 124884;
Best Local Similarity 100.0%; Pred. No. 1.3e-204;
Matches 705; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ATGACACAACCCGCATCGTCTCGTGTAGTCTTTGATCCCAGCAACCCCACCACATTTTCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      43138 ATGACACAACCCGCATCGTCTCGTGTAGTCTTTGATCCCAGCAACCCCACCACATTTTCG 43079

Qy         61 GTGGAAGCAATTGCGGCTTACACCCCCGTTGCTTTAATACGACTTTTAAACGCCAGTGGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      43078 GTGGAAGCAATTGCGGCTTACACCCCCGTTGCTTTAATACGACTTTTAAACGCCAGTGGA 43019

Qy        121 CCTTTGCAACCTGGTCACCGTGTGGACATCGCTGATGCCAGAAGCATTTACACCGTGGGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      43018 CCTTTGCAACCTGGTCACCGTGTGGACATCGCTGATGCCAGAAGCATTTACACCGTGGGA 42959

Qy        181 GCCGCGGCCAGTGCCGCGCGTGCACGCGCTAACCATAATGCAAATACGATACGCCGAACG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42958 GCCGCGGCCAGTGCCGCGCGTGCACGCGCTAACCATAATGCAAATACGATACGCCGAACG 42899

Qy        241 GCCATGTTTGCCGAGACTGACCCTATGACATGGTTAAGACCAACGGTTGGCTTAAAACGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42898 GCCATGTTTGCCGAGACTGACCCTATGACATGGTTAAGACCAACGGTTGGCTTAAAACGT 42839

Qy        301 ACGTTTAACCCGCGTATTATACGACCACAACCCCCAAATCCATCCATGAGTTTGGGAATC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42838 ACGTTTAACCCGCGTATTATACGACCACAACCCCCAAATCCATCCATGAGTTTGGGAATC 42779

Qy        361 TCGGGGCCTACTATATTGCCGCAAAAAACACAGAGCGCCGATCAGTCTGCTTTACAACAG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db      42778 TCGGGGCCTACTATATTGCCGCAAAAAACACAGAGCGCCGATCAGTCTGCTTTACAACAG 42719

Qy        421 CCCGCCGCGTTGGCGTTTTCGGGATCATCCCCGCAACACCCCCCACCTCAAACGACGTCG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42718 CCCGCCGCGTTGGCGTTTTCGGGATCATCCCCGCAACACCCCCCACCTCAAACGACGTCG 42659

Qy        481 GCATCCGTTGGACAACAGCAACACGTGGTGTCGGGGTCTTCTGGACAACAACCGCAACAG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42658 GCATCCGTTGGACAACAGCAACACGTGGTGTCGGGGTCTTCTGGACAACAACCGCAACAG 42599

Qy        541 GGAGCACAGTCAAGCACTGTCCAGCCAACAACCGGATCACCGCCCGCGGCCCAAGGCGTG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42598 GGAGCACAGTCAAGCACTGTCCAGCCAACAACCGGATCACCGCCCGCGGCCCAAGGCGTG 42539

Qy        601 CCACAGTCTACCCCGCCCCCAACCCAAAATACCCCCCAGGGGGGTAAGGGACAGACCTTG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42538 CCACAGTCTACCCCGCCCCCAACCCAAAATACCCCCCAGGGGGGTAAGGGACAGACCTTG 42479

Qy        661 TCACACACGGGACAATCTGGAAACGCTTCAAGAAGTCGTAGGGTG 705
              |||||||||||||||||||||||||||||||||||||||||||||
Db      42478 TCACACACGGGACAATCTGGAAACGCTTCAAGAAGTCGTAGGGTG 42434
```

RESULT 6
AR431491/c
LOCUS AR431491 124884 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 1 from patent US 6653069.
ACCESSION AR431491
VERSION AR431491.1 GI:40193595
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 124884)
AUTHORS Gomi,Y., Sunamachi,H., Takahashi,M. and Yamanishi,K.
TITLE Method for quality control of an attenuated vericella live vaccine
JOURNAL Patent: US 6653069-A 1 25-NOV-2003;
FEATURES Location/Qualifiers
source 1. .124884
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN

Query Match 100.0%; Score 705; DB 6; Length 124884;
Best Local Similarity 100.0%; Pred. No. 1.3e-204;
Matches 705; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ATGACACAACCCGCATCGTCTCGTGTAGTCTTTGATCCCAGCAACCCCACCACATTTTCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      43138 ATGACACAACCCGCATCGTCTCGTGTAGTCTTTGATCCCAGCAACCCCACCACATTTTCG 43079

Qy         61 GTGGAAGCAATTGCGGCTTACACCCCCGTTGCTTTAATACGACTTTTAAACGCCAGTGGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      43078 GTGGAAGCAATTGCGGCTTACACCCCCGTTGCTTTAATACGACTTTTAAACGCCAGTGGA 43019

Qy        121 CCTTTGCAACCTGGTCACCGTGTGGACATCGCTGATGCCAGAAGCATTTACACCGTGGGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      43018 CCTTTGCAACCTGGTCACCGTGTGGACATCGCTGATGCCAGAAGCATTTACACCGTGGGA 42959

Qy        181 GCCGCGGCCAGTGCCGCGCGTGCACGCGCTAACCATAATGCAAATACGATACGCCGAACG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42958 GCCGCGGCCAGTGCCGCGCGTGCACGCGCTAACCATAATGCAAATACGATACGCCGAACG 42899

Qy        241 GCCATGTTTGCCGAGACTGACCCTATGACATGGTTAAGACCAACGGTTGGCTTAAAACGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42898 GCCATGTTTGCCGAGACTGACCCTATGACATGGTTAAGACCAACGGTTGGCTTAAAACGT 42839

Qy        301 ACGTTTAACCCGCGTATTATACGACCACAACCCCCAAATCCATCCATGAGTTTGGGAATC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42838 ACGTTTAACCCGCGTATTATACGACCACAACCCCCAAATCCATCCATGAGTTTGGGAATC 42779

Qy        361 TCGGGGCCTACTATATTGCCGCAAAAAACACAGAGCGCCGATCAGTCTGCTTTACAACAG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db      42778 TCGGGGCCTACTATATTGCCGCAAAAAACACAGAGCGCCGATCAGTCTGCTTTACAACAG 42719

Qy        421 CCCGCCGCGTTGGCGTTTTCGGGATCATCCCCGCAACACCCCCCACCTCAAACGACGTCG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42718 CCCGCCGCGTTGGCGTTTTCGGGATCATCCCCGCAACACCCCCCACCTCAAACGACGTCG 42659

Qy        481 GCATCCGTTGGACAACAGCAACACGTGGTGTCGGGGTCTTCTGGACAACAACCGCAACAG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42658 GCATCCGTTGGACAACAGCAACACGTGGTGTCGGGGTCTTCTGGACAACAACCGCAACAG 42599

Qy        541 GGAGCACAGTCAAGCACTGTCCAGCCAACAACCGGATCACCGCCCGCGGCCCAAGGCGTG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42598 GGAGCACAGTCAAGCACTGTCCAGCCAACAACCGGATCACCGCCCGCGGCCCAAGGCGTG 42539

Qy        601 CCACAGTCTACCCCGCCCCCAACCCAAAATACCCCCCAGGGGGGTAAGGGACAGACCTTG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42538 CCACAGTCTACCCCGCCCCCAACCCAAAATACCCCCCAGGGGGGTAAGGGACAGACCTTG 42479

Qy        661 TCACACACGGGACAATCTGGAAACGCTTCAAGAAGTCGTAGGGTG 705
              |||||||||||||||||||||||||||||||||||||||||||||
Db      42478 TCACACACGGGACAATCTGGAAACGCTTCAAGAAGTCGTAGGGTG 42434


RESULT 7
BD097170/c
LOCUS       BD097170               124884 bp    DNA     linear   PAT 27-AUG-2002
DEFINITION  Method for the quality control of attetunated varicella live
            vaccine.
ACCESSION   BD097170
VERSION     BD097170.1  GI:22642744
KEYWORDS    WO 0156600-A/1.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 124884)
  AUTHORS   Gomi,Y., Sunamachi,H., Takahashi,M. and Yamanishi,K.
  TITLE     Method for the quality control of attetunated varicella live
            vaccine
  JOURNAL   Patent: WO 0156600-A 1 09-AUG-2001;
            THE RESEARCH FOUNDATION FOR MICROBIAL DISEASES OF OSAKA UNIVERSITY,
            YASUYUKI GOMI,HIROKI SUNAMACHI,MICHIAKI TAKAHASHI,KOICHI YAMANISHI
COMMENT     OS   Varicella virus Dumas strain
            PN   WO 0156600-A/1
            PD   09-AUG-2001
            PF   31-JAN-2001 WO 2001JP000678
            PR   31-JAN-2000 JP 00P    62734
            PI   YASUYUKI GOMI,HIROKI SUNAMACHI,MICHIAKI TAKAHASHI,KOICHI  PI
            YAMANISHI
            PC   A61K39/25,A61P31/22,C12N15/38,C12N7/00,G01N33/15 CC   Method
            for the quality control of attetunated varicella live  CC
             vaccine
            FH   Key             Location/Qualifiers
            FT   source          1. .124884
            FT                   /organism='Varicella virus Dumas strain'.
FEATURES             Location/Qualifiers
     source          1. .124884
                     /organism="unidentified"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32644"
ORIGIN

  Query Match             100.0%;  Score 705;  DB 6;  Length 124884;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-204;
  Matches  705;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGACACAACCCGCATCGTCTCGTGTAGTCTTTGATCCCAGCAACCCCACCACATTTTCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      43138 ATGACACAACCCGCATCGTCTCGTGTAGTCTTTGATCCCAGCAACCCCACCACATTTTCG 43079

Qy         61 GTGGAAGCAATTGCGGCTTACACCCCCGTTGCTTTAATACGACTTTTAAACGCCAGTGGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      43078 GTGGAAGCAATTGCGGCTTACACCCCCGTTGCTTTAATACGACTTTTAAACGCCAGTGGA 43019

Qy        121 CCTTTGCAACCTGGTCACCGTGTGGACATCGCTGATGCCAGAAGCATTTACACCGTGGGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      43018 CCTTTGCAACCTGGTCACCGTGTGGACATCGCTGATGCCAGAAGCATTTACACCGTGGGA 42959

Qy        181 GCCGCGGCCAGTGCCGCGCGTGCACGCGCTAACCATAATGCAAATACGATACGCCGAACG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42958 GCCGCGGCCAGTGCCGCGCGTGCACGCGCTAACCATAATGCAAATACGATACGCCGAACG 42899

Qy        241 GCCATGTTTGCCGAGACTGACCCTATGACATGGTTAAGACCAACGGTTGGCTTAAAACGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42898 GCCATGTTTGCCGAGACTGACCCTATGACATGGTTAAGACCAACGGTTGGCTTAAAACGT 42839

Qy        301 ACGTTTAACCCGCGTATTATACGACCACAACCCCCAAATCCATCCATGAGTTTGGGAATC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42838 ACGTTTAACCCGCGTATTATACGACCACAACCCCCAAATCCATCCATGAGTTTGGGAATC 42779

Qy        361 TCGGGGCCTACTATATTGCCGCAAAAAACACAGAGCGCCGATCAGTCTGCTTTACAACAG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42778 TCGGGGCCTACTATATTGCCGCAAAAAACACAGAGCGCCGATCAGTCTGCTTTACAACAG 42719

Qy        421 CCCGCCGCGTTGGCGTTTTCGGGATCATCCCCGCAACACCCCCCACCTCAAACGACGTCG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42718 CCCGCCGCGTTGGCGTTTTCGGGATCATCCCCGCAACACCCCCCACCTCAAACGACGTCG 42659

Qy        481 GCATCCGTTGGACAACAGCAACACGTGGTGTCGGGGTCTTCTGGACAACAACCGCAACAG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42658 GCATCCGTTGGACAACAGCAACACGTGGTGTCGGGGTCTTCTGGACAACAACCGCAACAG 42599

Qy        541 GGAGCACAGTCAAGCACTGTCCAGCCAACAACCGGATCACCGCCCGCGGCCCAAGGCGTG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42598 GGAGCACAGTCAAGCACTGTCCAGCCAACAACCGGATCACCGCCCGCGGCCCAAGGCGTG 42539

Qy        601 CCACAGTCTACCCCGCCCCCAACCCAAAATACCCCCCAGGGGGGTAAGGGACAGACCTTG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42538 CCACAGTCTACCCCGCCCCCAACCCAAAATACCCCCCAGGGGGGTAAGGGACAGACCTTG 42479

Qy        661 TCACACACGGGACAATCTGGAAACGCTTCAAGAAGTCGTAGGGTG 705
              |||||||||||||||||||||||||||||||||||||||||||||
Db      42478 TCACACACGGGACAATCTGGAAACGCTTCAAGAAGTCGTAGGGTG 42434


RESULT 8
HEVZVXX/c
LOCUS       HEVZVXX                124884 bp    DNA     linear   VRL 26-AUG-2004
DEFINITION  Human herpesvirus 3 (strain Dumas) complete genome.
ACCESSION   X04370 M14891 M16612
VERSION     X04370.1  GI:59989
KEYWORDS    ATPase subunit of terminase; complete genome; deoxypyrimidine
            kinase; deoxyribonuclease; deoxyuridine triphosphatase; DNA
            polymerase catalytic subunit; envelope protein; glycoprotein;
            helicase; host shut-off factor; major capsid protein; major capsid
            scaffold protein; minor capsid protein; minor capsid protein
            binding protein; nuclear matrix protein; nuclear
            membrane-associated phosphoprotein; portal protein; primase;
            processivity factor; protease; ribonucleotide reductase;
            ring-finger protein; serine-threonine protein kinase;
            single-stranded DNA-binding protein; small capsid protein;
            terminase subunit; thymidylate synthase; transcription regulator;
            uracil-DNA glycosylase.
SOURCE      Human herpesvirus 3 (strain Dumas)
  ORGANISM  Human herpesvirus 3 (strain Dumas)
            Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
            Alphaherpesvirinae; Varicellovirus.
REFERENCE   1  (bases 1 to 124884)
  AUTHORS   Davison,A.J. and Scott,J.E.
  TITLE     The complete DNA sequence of varicella-zoster virus
  JOURNAL   J. Gen. Virol. 67 (Pt 9), 1759-1816 (1986)
  MEDLINE   86306657
   PUBMED   3018124
FEATURES             Location/Qualifiers
     source          1. .124884
                     /organism="Human herpesvirus 3 (strain Dumas)"
                     /mol_type="genomic DNA"
```

```
                /strain="Dumas"
                /db_xref="taxon:10338"
                /note="acronym:HHV-3=VZV"
misc_feature    88. .89
                /note="TRL/UL boundary"
CDS             173. .562
                /function="putative transmembrane protein"
                /note="unnamed protein product; ORF0
                orthologue of HHV-1 UL56"
                /codon_start=1
                /protein_id="CAG46438.1"
                /db_xref="GI:49457166"
                /db_xref="UniProt/TrEMBL:Q6F6K2"
                /translation="MATVHYSRRPGTPPVTLTSSPSMDDVATPIPYLPTYAEAVADAP
                PPYRSRESLVFSPPLFPHVENGTTQQSYDCLDCAYDGIHRLQLAFLRIRKCCVPAFLI
                LFGILTLTAVVVAIVAVFPEEPPNSTT"
polyA_signal    complement(588. .593)
                /note="ORF1"
CDS             complement(589. .915)
                /function="putative transmembrane protein"
                /note="unnamed protein product; ORF1"
                /codon_start=1
                /protein_id="CAA27884.1"
                /db_xref="GI:59990"
                /db_xref="UniProt/Swiss-Prot:P09266"
                /translation="MSRVSEYGVPEGVRESDSDTDSVFMYQHTELMQNNASPLVVQTR
                PPAVLIPLVDVPRPRSRRKASAQLKMQMDRLCNVLGVVLQMATLALVTYIAFVVHTRA
                TSCKRE"
CDS             1134. .1850
                /note="unnamed protein product; ORF2"
                /codon_start=1
                /protein_id="CAA27885.1"
                /db_xref="GI:59991"
                /db_xref="UniProt/Swiss-Prot:P09267"
                /translation="MHVISETLAYGHVPAFIMGSTLVRPSLNATAEENPASETRCLLR
                VLAGRTVDLPGGGTLHITCTKTYVIIGKYSKPGERLSLARLIGRAMTPGGARTFIILA
                MKEKRSTTLGYECGTGLHLLAPSMGTFLRTHGLSNRDLCLWRGNIYDMHMQRLMFWEN
                IAQNTTETPCITSTLTCNLTEDSGEAALTTSDRPTLPTLTAQGRPTVSNIRGILKGSP
                RQQPVCHRVRFAEPTEGVLM"
polyA_signal    1857. .1862
                /note="ORF2"
polyA_signal    complement(1890. .1895)
                /note="ORF3"
CDS             complement(1908. .2447)
                /note="unnamed protein product; ORF3
                orthologue of HHV-1 UL55"
                /codon_start=1
                /protein_id="CAA27886.1"
                /db_xref="GI:59992"
                /db_xref="GOA:P09268"
                /db_xref="UniProt/Swiss-Prot:P09268"
                /translation="MDTTGASESSQPIRVNLKPDPLASFTQVIPPLALETTWTCPANS
                HAPTPSPLYGVKRLCALRATCGRADDLHAFLIGLGRRDKPSESPMYVDLQPFCSLLNS
                QRLLPEMANYNTLCDAPFSAATQQMMLESGQLGVHLAAIGYHCHCKSPFSAECWTGAS
                EAYDHVVCGGKARAAVGGL"
polyA_signal    complement(2777. .2782)
                /note="ORF4"
CDS             complement(2783. .4141)
                /function="post-transcriptional regulator of gene
                expression"
                /note="unnamed protein product; ORF4
                orthologue of HHV-1 UL54"
                /codon_start=1
                /protein_id="CAA27887.1"
                /db_xref="GI:59993"
                /db_xref="GOA:P09269"
                /db_xref="UniProt/Swiss-Prot:P09269"
                /translation="MASASIPTDPDVSTICEDFMNLLPDEPSDDFALEVTDWANDEAI
                GSTPGEDSTTSRTVYVERTADTAYNPRYSKRRHGRRESYHHNRPKTLVVVLPDSNHHG
                GRDVETGYARIERGHRRSSRSYNTQSSRKHRDRSLSNRRRRPTTPPAMTTGERNDQTH
                DESYRLRFSKRDARRERIRKEYDIPVDRITGRAIEVVSTAGASVTIDSVRHLDETIEK
                LVVRYATIQEGDSWASGGCFPGIKQNTSWPELMLYGHELYRTFESYKMDSRIARALRE
                RVIRGESLIEALESADELLTWIKMLAAKNLPIYTNNPIVATSKSLLENLKLKLGPFVR
                CLLLNRDNDLGSRTLPELLRQQRFSDITCITTYMFVMIARIANIVVRGSKFVEYDDIS
                CNVQVLQEYTPGSCLAGVLEALITHQRECGRVECTLSTWAGHLSDARPYGKYFKCSTF
                NC"
CDS             complement(4252. .5274)
                /note="ORF5
                orthologue of HHV-1 UL53"
                /codon_start=1
                /product="virion glycoprotein gK"
                /protein_id="CAA27888.1"
                /db_xref="GI:59994"
                /db_xref="GOA:P09261"
                /db_xref="UniProt/Swiss-Prot:P09261"
                /translation="MQALGIKTEHFIIMCLLSGHAVFTLWYTARVKFEHECVYATTVI
                NGGPVVWGSYNNSLIYVTFVNHSTFLDGLSGYDYSCRENLLSGDTMVKTAISTPLHDK
                IRIVLGTRNCHAYFWCVQLKMIFFAWFVYGMYLQFRRIRRMFGPFRSSCELISPTSYS
                LNYVTRVISNILLGYPYTKLARLLCDVSMRRDGMSKVFNADPISFLYMHKGVTLLMLL
                EVIAHISSGCIVLLTLGVAYTPCALLYPTYIRILAWVVVCTLAIVELISYVRPKPTKD
                NHLNHINTGGIRGICTTCCATVMSGLAIKCFYIVIFAIAVVIFMHYEQRVQVSLFGES
                ENSQKH"
CDS             complement(5326. .8577)
                /function="component of DNA helicase-primase complex"
                /note="ORF6
                orthologue of HHV-1 UL52"
                /codon_start=1
                /product="primase"
                /protein_id="CAA27889.1"
                /db_xref="GI:59995"
                /db_xref="GOA:P09270"
                /db_xref="UniProt/Swiss-Prot:P09270"
                /translation="MDKSSKPTIRLLFATKGCAISHSLLLLTGQISTEPLYVVSYTWT
                PDLDDVFVKNGREEITQVIPTKRPREVTENDEENQIMHLFCSRDVNVIFYLIGGFSTG
                DVRSRVWPIFFCCFKTQTDFKALYKALWYGAPLNPHIISDTLCISETFDIHSEVIQTL
                MVTTHHLNRKGLSDNGLCITEATLCKLVKKSVGRQELTSLYAHYERQVLAAYRRLYWG
                YGCSPFWYIVRFGPSEKTLVLATRYYLLQTDTSYNTLETPLYDLQAIKDLFLTYQVPA
                LPNCSGYNISDLLSFDKLSMFCCSSTYTRGLTAKNALSYILQRIHTDTTEIHAVSEYI
                TNDRKGLKVPDREFVDYIYLAHFECFNRKQIADHLQAVTYSDFVNKPVLLKSSNLGKR
                ATANFFNHVRSRLNMRDYIKKNVICDVTELGPEIGHKYTITKTYTLSLTYAAKPSKFI
                GVCDLATTLTRRVENIEKQFSPYGWSSTIPSNPPGFDELSNFEDSGVSAEALRAANFA
                NDTPNQSGRTGFDTSPGITKLLLFFSAATGIATHDVSILSYKTPLEALIGHSEVTGPM
                PVYRVALPHGAQAFAVIANDTWSSITNRYTLPHEARLIAEDLKQINPCNFVAASLRDM
                QLTLLLSTSVKNVSKISSNIPKDQLYINRNELFNTNLIITNLILDVDFHIRKPIPLGI
                LHAGMRAFRHGILTAMQLLFPKAVVNPNKDPCYFYKTACPEPTVEVLDDDNLLDITSH
                SDIDFYIENGELYTCVEENYTEDVWFFDTQTTSEVHTHADVSNNENLHETLPCNCKEK
                IGFRVCVPIPNPYALVGSSTLKGFAQILQQAVLLEREFVEYIGPYLRDFSFIDTGVYS
                HGHSLRLPFFSKVTTTGTAVGQLLPFYVVPEQCIDILAFVTSHRNPANFHFHSRPQSN
                VPVQFILHNLGGEYAEFFERKVARNKQIFSSPQISLTKALKERGVTCLDAFTLEAFVD
                STILESIVEHIAVHFPGRDREYTLTSSKCIAIKRDWVLFQLICGTKGFTCLRYPHRGG
                RTAPRTFVSLRVDHHNRLCISLAQQCFATKCDSNRMHTIFTLEVPNYPNLTSS"
CDS             8607. .9386
                /function="tegument protein"
                /note="unnamed protein product; ORF7
                orthologue of HHV-1 UL51"
                /codon_start=1
                /protein_id="CAA27890.1"
                /db_xref="GI:59996"
                /db_xref="UniProt/Swiss-Prot:P09271"
                /translation="MQTVCASLCGYARIPTEEPSYEEVRVNTHPQGAALLRLQEALTA
                VNGLLPAPLTLEDVVASADNTRRLVRAQALARTYAACSRNIECLKQHHFTEDNPGLNA
                VVRSHMENSKRLADMCLAAITHLYLSVGAVDVTTDDIVDQTLRMTAESEVVMSDVVLL
                EKTLGVVAKPQASFDVSHNHELSIAKGENVGLKTSPIKSEATQLSEIKPPLIEVSDNN
                TSNLTKKTYPTETLQPVLTPKQTQDVQRTTPAIKKSHVMLV"
polyA_signal    9394. .9399
                /note="ORF7"
polyA_signal    complement(9426. .9431)
                /note="ORF8"
CDS             complement(9477. .10667)
                /note="ORF8
                orthologue of HHV-1 UL50"
                /codon_start=1
                /product="deoxyuridine triphosphatase"
                /protein_id="CAA27891.1"
                /db_xref="GI:59997"
                /db_xref="GOA:P09254"
```

```
                     /db_xref="UniProt/Swiss-Prot:P09254"
                     /translation="MNEAVIDPILETAVNTGDMFCSQTIPNRCLKDTILIEVQPECAD
                     TLQCVLDDKVSRHQPLLLRNHKKLELPSEKSVTRGGFYMQQLELLVKSAPPNEYALLL
                     IQCKDTALADEDNFFVANGVIDAGYRGVISALLYYRPGVTVILPGHLTIYLFPVKLRQ
                     SRLLPKNVLKHLDPIFKSIQVQPLSNSPSNYEKPVIPEFADISTVQQGQPLHRDSAEY
                     HIDVPLTYKHIINPKRQEDAGYDICVPYNLYLKRNEFIKIVLPIIRDWDLQHPSINAY
                     IFGRSSKSRSGIIVCPTAWPAGEHCKFYVYNLTGDDIRIKTGDRLAQVLLIDHNTQIH
                     LKHNVLSNIAFPYAIRGKCGIPGVQWYFTKTLDLIATPSERGTRGFGSTDKETNDVDF
                     LLKH"
     CDS             10642. .10905
                     /function="disulphide-linked to gM"
                     /note="ORF9A
                     orthologue of HHV-1 UL49A"
                     /codon_start=1
                     /product="envelope protein gN"
                     /protein_id="CAH19068.1"
                     /db_xref="GI:51571542"
                     /translation="MGSITASFILITMQILFFCEDSSGEPNFAERNFWHASCSARGVY
                     IDGSMITTLFFYASLLGVCVALISLAYHACFRLFTRSVLRSTW"
     CDS             11009. .11917
                     /function="tegument protein"
                     /note="unnamed protein product; ORF9
                     orthologue of HHV-1 UL49"
                     /codon_start=1

  Query Match             100.0%;  Score 705;  DB 14;  Length 124884;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-204;
  Matches  705;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGACACAACCCGCATCGTCTCGTGTAGTCTTTGATCCCAGCAACCCCACCACATTTTCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      43138 ATGACACAACCCGCATCGTCTCGTGTAGTCTTTGATCCCAGCAACCCCACCACATTTTCG 43079

Qy         61 GTGGAAGCAATTGCGGCTTACACCCCCGTTGCTTTAATACGACTTTTAAACGCCAGTGGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      43078 GTGGAAGCAATTGCGGCTTACACCCCCGTTGCTTTAATACGACTTTTAAACGCCAGTGGA 43019

Qy        121 CCTTTGCAACCTGGTCACCGTGTGGACATCGCTGATGCCAGAAGCATTTACACCGTGGGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      43018 CCTTTGCAACCTGGTCACCGTGTGGACATCGCTGATGCCAGAAGCATTTACACCGTGGGA 42959

Qy        181 GCCGCGGCCAGTGCCGCGCGTGCACGCGCTAACCATAATGCAAATACGATACGCCGAACG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42958 GCCGCGGCCAGTGCCGCGCGTGCACGCGCTAACCATAATGCAAATACGATACGCCGAACG 42899

Qy        241 GCCATGTTTGCCGAGACTGACCCTATGACATGGTTAAGACCAACGGTTGGCTTAAAACGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42898 GCCATGTTTGCCGAGACTGACCCTATGACATGGTTAAGACCAACGGTTGGCTTAAAACGT 42839

Qy        301 ACGTTTAACCCGCGTATTATACGACCACAACCCCCAAATCCATCCATGAGTTTGGGAATC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42838 ACGTTTAACCCGCGTATTATACGACCACAACCCCCAAATCCATCCATGAGTTTGGGAATC 42779

Qy        361 TCGGGGCCTACTATATTGCCGCAAAAAACACAGAGCGCCGATCAGTCTGCTTTACAACAG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42778 TCGGGGCCTACTATATTGCCGCAAAAAACACAGAGCGCCGATCAGTCTGCTTTACAACAG 42719

Qy        421 CCCGCCGCGTTGGCGTTTTCGGGATCATCCCCGCAACACCCCCCACCTCAAACGACGTCG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42718 CCCGCCGCGTTGGCGTTTTCGGGATCATCCCCGCAACACCCCCCACCTCAAACGACGTCG 42659

Qy        481 GCATCCGTTGGACAACAGCAACACGTGGTGTCGGGGTCTTCTGGACAACAACCGCAACAG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42658 GCATCCGTTGGACAACAGCAACACGTGGTGTCGGGGTCTTCTGGACAACAACCGCAACAG 42599

Qy        541 GGAGCACAGTCAAGCACTGTCCAGCCAACAACCGGATCACCGCCCGCGGCCCAAGGCGTG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42598 GGAGCACAGTCAAGCACTGTCCAGCCAACAACCGGATCACCGCCCGCGGCCCAAGGCGTG 42539

Qy        601 CCACAGTCTACCCCGCCCCCAACCCAAAATACCCCCCAGGGGGGTAAGGGACAGACCTTG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42538 CCACAGTCTACCCCGCCCCCAACCCAAAATACCCCCCAGGGGGGTAAGGGACAGACCTTG 42479

Qy        661 TCACACACGGGACAATCTGGAAACGCTTCAAGAAGTCGTAGGGTG 705
              |||||||||||||||||||||||||||||||||||||||||||||
Db      42478 TCACACACGGGACAATCTGGAAACGCTTCAAGAAGTCGTAGGGTG 42434


RESULT 9
AY548171/c
LOCUS       AY548171               125459 bp    DNA     linear   VRL 14-JUN-2004
DEFINITION  Human herpesvirus 3 strain BC, complete genome.
ACCESSION   AY548171
VERSION     AY548171.1  GI:46981409
KEYWORDS    .
SOURCE      Human herpesvirus 3
  ORGANISM  Human herpesvirus 3
            Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
            Alphaherpesvirinae; Varicellovirus.
REFERENCE   1  (bases 1 to 125459)
  AUTHORS   Grose,C., Tyler,S., Peters,G., Hiebert,J., Stephens,G.M.,
            Ruyechan,W.T., Jackson,W., Storlie,J. and Tipples,G.A.
  TITLE     Complete DNA Sequence Analyses of the First Two Varicella-Zoster
            Virus Glycoprotein E (D150N) Mutant Viruses Found in North America:
            Evolution of Genotypes with an Accelerated Cell Spread Phenotype
  JOURNAL   J. Virol. 78 (13), 6799-6807 (2004)
   PUBMED   15194755
REFERENCE   2  (bases 1 to 125459)
  AUTHORS   Grose,C., Tyler,S., Peters,G., Hiebert,J., Stephens,G.,
            Ruyechan,W., Jackson,W., Storlie,J. and Tipples,G.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-FEB-2004) Health Canada, National Microbiology
            Laboratory, 1015 Arlington Street, Winnipeg, Manitoba R3E 3P6,
            Canada
FEATURES             Location/Qualifiers
     source          1. .125459
                     /organism="Human herpesvirus 3"
                     /mol_type="genomic DNA"
                     /strain="BC"
                     /db_xref="taxon:10335"
     misc_feature    88. .89
                     /note="TRL/UL boundary"
     CDS             239. .562
                     /note="HHV3g45"
                     /codon_start=1
                     /evidence=not_experimental
                     /product="hypothetical protein"
                     /protein_id="AAT07681.1"
                     /db_xref="GI:46981410"
                     /translation="MDDVATPIPYLPTYAEAVADAPPPYRSRESLVFSPPLFPHVENG
                     TTQQSYDCLDCAYDGIHRLQLAFLRIRKCCVPAFLILFGILTLTAVVVAIVAVFPEEP
                     PNSTT"
     CDS             complement(589. .915)
                     /note="hypothetical protein; HHV3g13"
                     /codon_start=1
                     /product="ORF1"
                     /protein_id="AAT07682.1"
                     /db_xref="GI:46981411"
                     /translation="MSRVSEYGVPEGVRESDSDTDSVFMYQHTELMQNNASPLVVQTR
                     PPAVLIPLVDVPRPRSRRKASAQLKMQMDRLCNVLGVVLQMATLALVTYIAFVVHTRA
                     TSCKRE"
     CDS             1134. .1850
                     /note="ORF2; HHV3g29"
                     /codon_start=1
                     /product="immediate-early protein"
                     /protein_id="AAT07683.1"
                     /db_xref="GI:46981412"
                     /translation="MHVISETLAYGHVPAFIMGSTLVRPSLNATAEENPASETRCLLR
                     VLAGRTVDLPGGGTLHITCTKTYVIIGKYSKPGERLSLARLIGRAMTPGGARTFIILA
                     MKEKRSTTLGYECGTGLHLLAPSMGTFLRTHGLSNRDLCLWRGNIYDMHMQRLMFWEN
                     IAQNTTETPCITSTLTCNLTEDSGEAALTTSDRPTLPTLTAQGRPTVSNIRGILKGSP
                     RQQPVCHRVRFAEPTEGVLM"
     CDS             complement(1908. .2447)
                     /note="ORF3; HHV3g30"
                     /codon_start=1
                     /evidence=not_experimental
```

/product="hypothetical protein"
/protein_id="AAT07684.1"
/db_xref="GI:46981413"
/translation="MDTTGASESSQPIRVNLKPDDPLASFTQVIPPLALETTWTCPANS
HAPTPSPLYGVKRLCALRATCGRADDLHAFLIGLGRRDKPSESPMYVDLQPFCSLLNS
QRLLPEMANYNTLCDAPFSAATQQMMLESGQLGVHLAAIGYHCHCKSPFSAECWTGAS
EAYDHVVCGGKARAAVGGL"

CDS complement(2783. .4141)
/note="ORF4; HHV3g21"
/codon_start=1
/product="transactivator"
/protein_id="AAT07685.1"
/db_xref="GI:46981414"
/translation="MASASIPTDPDVSTICEDFMNLLPDEPSDDFALEVTDWANDEAI
GSTPGEDSTTSRTVYVERTADTAYNPRYSKRRHGRRESYHHNRPKTLVVVLPDSNHHG
GRDVETGYARIERGHRRSSRSYNTQSSRKHRDRSLSNRRRRPTTPPAMTTGERNDQTH
DESYRLRFSKRDARRERIRKEYDIPVDRITGRAIEVVSTAGASVTIDSVRHLDETIEK
LVVRYATIQEGDSWASGGCFPGIKQNTSWPELMLYGHELYRTFESYKMDSRIARALRE
RVIRGESLIEALESADELLTWIKMLAAKNLPIYTNNPIVATSKSLLENLKLKLGPFVR
CLLLNRDNDLGSRTLPELLRQQRPSDITCITTYMFVMIARIANIVVRGSKFVEYDDIS
CNVQVLQEYTPGSCLAGVLEALITHQRECGRVECTLSTWAGHLSDARPYGKYFKCSTF
NC"

CDS complement(4252. .5274)
/note="ORF5; HHV3g22"
/codon_start=1
/product="glycoprotein K"
/protein_id="AAT07686.1"
/db_xref="GI:46981415"
/translation="MQALGIKTEHFIIMCLLSGHAVFTLWYTARVKFEHECVYATTVI
NGGPVVWGSYNNSLIYVTFVNHSTFLDGLSGYDYSCRENLLSGDTMVKTAISTPLHDK
IRIVLGTRNCHAYFWCVQLKMIFFAWFVYGMYLQFRRIRRMFGPFRSSCELISPTSYS
LNYVTRVISNILLGYPYTKLARLLCDVSMRRDGMSKVFNADPISFLYMHKGVTLLMLL
EVIAHISSGCIVLLTLGVAYTPCALLYPTYIRILAWVVVCTLAIVELISYVRPKPTKD
NHLNHINTGGIRGICTTCCATVMSGLAIKCFYIVIFAIAVVIFMHYEQRVQVSLFGES
ENSQKH"

CDS complement(5326. .8577)
/note="ORF6; HHV3g25"
/codon_start=1
/product="DNA helicase-primase complex"
/protein_id="AAT07687.1"
/db_xref="GI:46981416"
/translation="MDKSSKPTIRLLFATKGCAISHSLLLLTGQISTEPLYVVSYTWT
PDLDDVFVKNGREEITQVIPTKRPREVTENDEENQIMHLFCSRDVNVIFYLIGGFSTG
DVRSRVWPIFFCCFKTQTDFKALYKALWYGAPLNPHIISDTLCISETFDIHSEVIQTL
MVTTHHLNRKGLSDNGLCITEATLCKLVKKSVGRQELTSLYAHYERQVLAAYRRLYWG
YGCSPFWYIVRFGPSEKTLVLATRYYLLQTDTSYNTLETPLYDLQAIKDLFLTYQVPA
LPNCSGYNISDLLSFDKLSMFCCSSTYTRGLTAKNALSYILQRIHTDTTEIHAVSEYI
TNDRKGLKVPDREFVDYIYLAHFECFNRKQIADHLQAVTYSDFVNKPVLLKSSNLGKR
ATANFFNHVRSRLNMRDYIKKNVICDVTELGPEIGHKYTITKTYTLSLTYAAKPSKFI
GVCDLATTLTRRVENIEKQFSPYGWSSTIPSNPPGFDELSNFEDSGVSAEALRAANFA
NDTPNQSGRTGFDTSPGITKLLLFFSAATGIATHDVSILSYKTPLEALIGHSEVTGPM
PVYRVALPHGAQAFAVIANDTWSSITNRYTLPHEARLIAEDLKQINPCNFVAASLRDM
QLTLLLSTSVKNVSKISSNIPKDQLYINRNELFNTNLIITNLILDVDFHIRKPIPLGI
LHAGMRAFRHGILTAMQLLFPKAVVNPNKDPCYFYKTACPEPTVEVLDDDNLLDITSH
SDIDFYIENGELYTCVEENYTEDVWFFDTQTTSEVHTHADVSNNENLHETLPCNCKEK
IGFRVCVPIPNPYALVGSSTLKGFAQILQQAVLLEREFVEYIGPYLRDFSFIDTGVYS
HGHSLRLPFFSKVTTTGTAVGQLLPFYVVPEQCIDILAFVTSHRNPANFHFHSRPQSN
VPVQFILHNLGGEYAEFFERKVARNKQIFSSPQISLTKALKERGVTCLDAFTLEAFVD
STILESIVEHIAVHFPGRDREYTLTSSKCIAIKRDWVLFQLICGTKGFTCLRYPHRGG
RTAPRTFVSLRVDHHNRLCISLAQQCFATKCDSNRMHTIFTLEVPNYPNLTSS"

CDS 8607. .9386
/note="ORF7; HHV3g26"
/codon_start=1
/product="virion phosphoprotein"
/protein_id="AAT07688.1"
/db_xref="GI:46981417"
/translation="MQTVCASLCGYARIPTEEPSYEEVRVNTHPQGAALLRLQEALTA
VNGLLPAPLTLEDVVASADNTRRLVRAQALARTYAACSRNIECLKQHHFTEDNPGLNA
VVRSHMENSKRLADMCLAAITHLYLSVGAVDVTTDDIVDQTLRMTAESEVVMSDVVLL
EKTLGVVAKPQASFDVSHNHELSIAKGENVGLKTSPIKSEATQLSEIKPPLIEVSDNN
TSNLTKKTYPTETLQPVLTPKQTQDVQRTTPAIKKSHVMLV"

CDS complement(9477. .10667)
/note="ORF8; HHV3g20"
/codon_start=1
/product="dUTPase"
/protein_id="AAT07689.1"
/db_xref="GI:46981418"
/translation="MNEAVIDPILETAVNTGDMFCSQTIPNRCLKDTILIEVQPECAD
TLQCVLDDKVSRHQPLLLRNHKKLELPSEKSVTRGGFYMQQLELLVKSAPPNEYALLL
IQCKDTALADEDNFFVANGVIDAGYRGVISALLYYRPGVTVILPGHLTIYLFPVKLRQ
SRLLPKNVLKHLDPIFKSIQVQPLSNSPSNYEKPVIPEFADISTVQQGQPLHRDSAEY
HIDVPLTYKHIINPKRQEDAGYDICVPYNLYLKRNEFIKIVLPIIRDWDLQHPSINAY
IFGRSSKSRSGIIVCPTAWPAGEHCKFYVYNLTGDDIRIKTGDRLAQVLLIDHNTQIH
LKHNVLSNIAFPYAIRGKCGIPGVQWYFTKTLDLIATPSERGTRGFGSTDKETNDVDF
LLKH"

CDS 10678. .10905
/note="Predicted; HHV3g46"
/codon_start=1
/product="membrane protein"
/protein_id="AAT07690.1"
/db_xref="GI:46981419"
/translation="MQILFFCEDSSGEPNFAERNFWHASCSARGVYIDGSMITTLFFY
ASLLGVCVALISLAYHACFRLFTRSVLRSTW"

CDS 11009. .11917
/note="ORF9; HHV3g23"
/codon_start=1
/product="tegument protein"
/protein_id="AAT07691.1"
/db_xref="GI:46981420"
/translation="MASSDGDRLCRSNAVRRKTTPSYSGQYRTARRSVVVGPPDDSDD
SLGYITTVGADSPSPVYADLYFEHKNTTPRVHQPNDSSGSEDDFEDIDEVVAAFREAR
LRHELVEDAVYENPLSVEKPSRSFTKNAAVKPKLEDSPKRAPPGAGAIASGRPISFST
APKTATSSWCGPTPSYNKRVFCEAVRRVAAMQAQKAAEAAWNSNPPRNNAELDRLLTG
AVIRITVHEGLNLIQAANEADLGEGASVSKRGHNRKTGDLQGGMGNEPMYAQVRKPKS
RTDTQTTGRITNRSRARSASRTDTRK"

CDS 12160. .13392
/note="ORF10; HHV3g24"
/codon_start=1
/product="transactivator"
/protein_id="AAT07692.1"
/db_xref="GI:46981421"
/translation="MECNLGTEHPSTDTWNRSKTEQAVVDAFDESLFGDVASDIGFET
SLYSHAVKTAPSPPWVASPKILYQQLIRDLDFSEGPRLLSCLETWNEDLFSCFPINED
LYSDMMVLSPDPDDVISTVSTKDHVEMFNLTTRGSVRLPSPPKQPTGLPAYVQEVQDS
FTVELRAREEAYTKLLVTYCKSIIRYLQGTAKRTTIGLNIQNPDQKAYTQLRQSILLR
YYREVASLARLLYLHLYLTVTREFSWRLYASQSAHPDVFAALKFTWTERRQFTCAFHP
VLCNHGIVLLEGKPLTASALREINYRRRELGLPLVRCGLVEENKSPLVQQPSFSVHLP
RSVGFLTHHIKRKLDAYAVKHPQEPRHVRADHPYAKVVENRNYGSSIEAMILAPPSPS
EILPGDPPRPPTCGFLTR"

CDS 13590. .16004
/note="ORF11; HHV3g03"
/codon_start=1
/product="tegument protein and transactivator"
/protein_id="AAT07693.1"
/db_xref="GI:46981422"
/translation="MQSGHYNRRQSRRQRISSNTTDSPRHTHGTRYRSTNWYTHPPQI
LSNSETLVAVQELLNSEMDQDSSSDASDDFPGYALHHSTYNGSEQNTSTSRHENRIFK
LTEREANEEININTDAIDDEGEAEEGEAEEDAIDDEGEAEEGEAEEDAIDDEGEAEEG
EAEEDAIDDEGEAEEDAAEEDAIDDEGEAEEDAIDDEGEAEEDYFSVSQVCSRDADEV
YFTLDPEISYSTDLRIAKVMEPAVSKELNVSKRCVEPVTLTGSMLAHNGFDESWFAMR
ECTRREYITVQGLYDPIHLRYQFDTSRMTPPQILRTIPALPNMTLGELLLIFPIEFMA
QPISIERILVEDVFLDRRASSKTHKYGPRWNSVYALPYNAGKMYVQHIPGFYDVSLRA
VGQGTAIWHHMILSTAACAISNRISHGDGLGFLLDAAIRISANCIFLGRNDNFGVGDP
CWLEDHLAGLPREAVPDVLQVTQLVLPNRGPTVAIMRGFFGALAYWPELRIAISEPST

Query Match 100.0%; Score 705; DB 14; Length 125459;
Best Local Similarity 100.0%; Pred. No. 1.3e-204;
Matches 705; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ATGACACAACCCGCATCGTCTCGTGTAGTCTTTGATCCCAGCAACCCCACCACATTTTCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      43632 ATGACACAACCCGCATCGTCTCGTGTAGTCTTTGATCCCAGCAACCCCACCACATTTTCG 43573

Qy         61 GTGGAAGCAATTGCGGCTTACACCCCCGTTGCTTTAATACGACTTTTAAACGCCAGTGGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      43572 GTGGAAGCAATTGCGGCTTACACCCCCGTTGCTTTAATACGACTTTTAAACGCCAGTGGA 43513
```

```
Qy        121 CCTTTGCAACCTGGTCACCGTGTGGACATCGCTGATGCCAGAAGCATTTACACCGTGGGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      43512 CCTTTGCAACCTGGTCACCGTGTGGACATCGCTGATGCCAGAAGCATTTACACCGTGGGA 43453

Qy        181 GCCGCGGCCAGTGCCGCGCGTGCACGCGCTAACCATAATGCAAATACGATACGCCGAACG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      43452 GCCGCGGCCAGTGCCGCGCGTGCACGCGCTAACCATAATGCAAATACGATACGCCGAACG 43393

Qy        241 GCCATGTTTGCCGAGACTGACCCTATGACATGGTTAAGACCAACGGTTGGCTTAAAACGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      43392 GCCATGTTTGCCGAGACTGACCCTATGACATGGTTAAGACCAACGGTTGGCTTAAAACGT 43333

Qy        301 ACGTTTAACCCGCGTATTATACGACCACAACCCCCAAATCCATCCATGAGTTTGGGAATC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      43332 ACGTTTAACCCGCGTATTATACGACCACAACCCCCAAATCCATCCATGAGTTTGGGAATC 43273

Qy        361 TCGGGGCCTACTATATTGCCGCAAAAAACACAGAGCGCCGATCAGTCTGCTTTACAACAG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      43272 TCGGGGCCTACTATATTGCCGCAAAAAACACAGAGCGCCGATCAGTCTGCTTTACAACAG 43213

Qy        421 CCCGCCGCGTTGGCGTTTTCGGGATCATCCCCGCAACACCCCCCACCTCAAACGACGTCG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      43212 CCCGCCGCGTTGGCGTTTTCGGGATCATCCCCGCAACACCCCCCACCTCAAACGACGTCG 43153

Qy        481 GCATCCGTTGGACAACAGCAACACGTGGTGTCGGGGTCTTCTGGACAACAACCGCAACAG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      43152 GCATCCGTTGGACAACAGCAACACGTGGTGTCGGGGTCTTCTGGACAACAACCGCAACAG 43093

Qy        541 GGAGCACAGTCAAGCACTGTCCAGCCAACAACCGGATCACCGCCCGCGGCCCAAGGCGTG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      43092 GGAGCACAGTCAAGCACTGTCCAGCCAACAACCGGATCACCGCCCGCGGCCCAAGGCGTG 43033

Qy        601 CCACAGTCTACCCCGCCCCCAACCCAAAATACCCCCCAGGGGGGTAAGGGACAGACCTTG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      43032 CCACAGTCTACCCCGCCCCCAACCCAAAATACCCCCCAGGGGGGTAAGGGACAGACCTTG 42973

Qy        661 TCACACACGGGACAATCTGGAAACGCTTCAAGAAGTCGTAGGGTG 705
              |||||||||||||||||||||||||||||||||||||||||||||
Db      42972 TCACACACGGGACAATCTGGAAACGCTTCAAGAAGTCGTAGGGTG 42928
```

```
RESULT 10
AB097932/c
LOCUS       AB097932               125078 bp    DNA     linear   VRL 13-DEC-2002
DEFINITION  Human herpesvirus 3 ORF62 gene for IE62 transactivator, complete
            cds, strain: The Oka vaccine.
ACCESSION   AB097932
VERSION     AB097932.1  GI:26665420
KEYWORDS    .
SOURCE      Human herpesvirus 3
  ORGANISM  Human herpesvirus 3
            Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
            Alphaherpesvirinae; Varicellovirus.
REFERENCE   1
  AUTHORS   Gomi,Y., Sunamachi,H., Mori,Y., Nagaike,K., Takahashi,M. and
            Yamanishi,K.
  TITLE     Comparison of the complete DNA sequences of the Oka varicella
            vaccine and its parental virus
  JOURNAL   J. Virol. 76 (22), 11447-11459 (2002)
  MEDLINE   22276345
   PUBMED   12388706
REFERENCE   2  (bases 1 to 125078)
  AUTHORS   Gomi,Y., Sunamachi,H., Nagaike,K., Mori,Y., Takahashi,M. and
            Yamanishi,K.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-DEC-2002) Yasuyuki Gomi, Kanonji institute, the
            research foundation for microbial diseases of Osaka university,
            research division; 2-9-41 Yahata-cho, Kanonji-city, Kagawa
            768-0062, Japan (E-mail:ygomi@mail.biken.or.jp,
            Tel:81-875-25-4171(ex.620), Fax:81-875-25-4843)
FEATURES             Location/Qualifiers
     source          1. .125078
                     /organism="Human herpesvirus 3"
                     /mol_type="genomic DNA"
                     /strain="The Oka vaccine"
                     /db_xref="taxon:10335"
                     /note="synonym: varicella-zoster virus"
     gene            complement(105189. .109121)
                     /gene="ORF62"
     CDS             complement(105189. .109121)
                     /gene="ORF62"
                     /codon_start=1
                     /product="IE62 transactivator"
                     /protein_id="BAC44881.1"
                     /db_xref="GI:26665421"
                     /translation="MDTPPMQRSTPQRAGSPDTLELMDLLDAAAAAAEHRARVVTSSQ
                     PDDLLFGENGVMVGREHEIVSIPSVSGLQPEPRTEDVGEELTQDDYVCEDGQDLXGSP
                     VIPLAEVFHTRFSEAGAREPTGADRSLETVSLGTKLARSPKPPMNDGETGRGTTPPFP
                     QAFSPVSPASPVGDAAGNDQREDQRSIPRQTTRGNSPGLPSVVHRDRQTQSISGKKPG
                     DEQAGHAHASGDGVVLQKTQRPAQGKSPKKKTLKVKVPLPARKPGGPVPGPVEQLYHV
                     LSDSVPAKGAKADLPFETDDTRPRKHDARGITPRVPGRSSGGKPRAFLALPGRSHAPD
                     PIEDDSPVEKKPKSREFVSSSSSSSSWGSSSEDEDDEPRRVSVGSETTGSRSGREHAP
                     SPSNSDDSDSNDGGSTKQNIQPGYRSISGPDPRIRKTKRLAGEPGRQRQKSFSLPRSR
                     TPIIPPVSGPLMMPDGSPWPGSAPLPSNRVRFGPSGETREGHWEDEAVRAARARYEAS
                     TEPXPLYVPELGDPARQYRALINLIYCPDRDPIAWLQNPKLTGVNSALNQFYQKLLPP
                     GRAGTAVTGSVASPVPHVGEAMATGEALWALPHAAAAVAMSRRYDRAQKHFILQSLRR
                     AFAGMAYPEATGSSPAARISRGHPSPTTPATQTPDPQPSAAARSLSVCPPDDRLRTPR
                     KRKSQPVESRSLLDKIRETPVADARVADDHVVSKAKRRVSEPVTITSGPVVDPPAVIT
                     MPLDGPAPNGGFRRIPRGALHTPVPSDQARKAYCTPETIARLVDDPLFPTAWRPALSF
                     DPGALAEIAARRPGGGDRRFGPPSGVEALRRRCAWMRQIPDPEDVRLLIIYDPLPGED
                     INGPLESTLATDPGPSWSPSRGGLSVVLAALSNRLCLPSTHAWAGNWTGPPDVSALNA
                     RGVLLLSTRDLAPAGAVEYLGSRLASARRRLLVLDAVALERWPGDGPALSQYHVYVRA
                     PARPDAQAVVRWPDSAVTEGLARAVFASSRTFGPASFARIETAFANLYPGEQPLCLCR
                     GGNVAYTVCTRAGPKTRVPLSPREYRQYVLPGFDGCKDLARQSRGLGLGAADFVDEAA
                     HSHRAANRWGLGAALRPVFLPEGRRPGAAGPEAGDVPTWARVFCRHALLEPDPAAEPL
                     VLPPVAGRSVALYASADEARNALPPIPRVMWPPGFGAAETVLEGSDGTRFAFGHHGGS
                     ERPAETQAGRQRRTADDREHALEPDDWEVGCEDAWDSEEGGGDDGDAPGSSFGVSVVS
                     VAPGVLRDRRVGXRPAVKVELLSSSSSSEDEDDVWGGRGGRSPPQSRG"
ORIGIN
```

```
  Query Match           99.8%;  Score 703.4;  DB 14;  Length 125078;
  Best Local Similarity   99.9%;  Pred. No. 4.2e-204;
  Matches  704;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 ATGACACAACCCGCATCGTCTCGTGTAGTCTTTGATCCCAGCAACCCCACCACATTTTCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      43041 ATGACACAACCCGCATCGTCTCGTGTAGTCTTTGATCCCAGCAACCCCACCACATTTTCG 42982

Qy         61 GTGGAAGCAATTGCGGCTTACACCCCCGTTGCTTTAATACGACTTTTAAACGCCAGTGGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42981 GTGGAAGCAATTGCGGCTTACACCCCCGTTGCTTTAATACGACTTTTAAACGCCAGTGGA 42922

Qy        121 CCTTTGCAACCTGGTCACCGTGTGGACATCGCTGATGCCAGAAGCATTTACACCGTGGGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42921 CCTTTGCAACCTGGTCACCGTGTGGACATCGCTGATGCCAGAAGCATTTACACCGTGGGA 42862

Qy        181 GCCGCGGCCAGTGCCGCGCGTGCACGCGCTAACCATAATGCAAATACGATACGCCGAACG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42861 GCCGCGGCCAGTGCCGCGCGTGCACGCGCTAACCATAATGCAAATACGATACGCCGAACG 42802

Qy        241 GCCATGTTTGCCGAGACTGACCCTATGACATGGTTAAGACCAACGGTTGGCTTAAAACGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42801 GCCATGTTTGCCGAGACTGACCCTATGACATGGTTAAGACCAACGGTTGGCTTAAAACGT 42742

Qy        301 ACGTTTAACCCGCGTATTATACGACCACAACCCCCAAATCCATCCATGAGTTTGGGAATC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42741 ACGTTTAACCCGCGTATTATACGACCACAACCCCCAAATCCATCCATGAGTTTGGGAATC 42682

Qy        361 TCGGGGCCTACTATATTGCCGCAAAAAACACAGAGCGCCGATCAGTCTGCTTTACAACAG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42681 TCGGGGCCTACTATATTGCCGCAAAAAACACAGAGCGCCGATCAGTCTGCTTTACAACAG 42622

Qy        421 CCCGCCGCGTTGGCGTTTTCGGGATCATCCCCGCAACACCCCCCACCTCAAACGACGTCG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42621 CCCGCCGCGTTGGCGTTTTCGGGATCATCCCCGCAACACCCCCCACCTCAAACGACGTCG 42562
```

```
Qy        481 GCATCCGTTGGACAACAGCAACACGTGGTGTCGGGGTCTTCTGGACAACAACCGCAACAG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42561 GCATCCGTTGGACAACAGCAACACGTGGTGTCGGGGTCTTCTGGACAACAACCGCAACAG 42502

Qy        541 GGAGCACAGTCAAGCACTGTCCAGCCAACAACCGGATCACCGCCCGCGGCCCAAGGCGTG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42501 GGAGCACAGTCAAGCACTGTCCAGCCAACAACCGGATCACCGCCCGCGGCCCAAGGCGTG 42442

Qy        601 CCACAGTCTACCCCGCCCCCAACCCAAAATACCCCCCAGGGGGGTAAGGGACAGACCTTG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42441 CCACAGTCTACCCCGCCCCCAACCCAAAATACCCCCCAGGGGGGTAAGGGACAGACCTTG 42382

Qy        661 TCACACACGGGACAATCTGGAAACGCTTCAAGAAGTCGTAGGGTG 705
              || |||||||||||||||||||||||||||||||||||||||||||
Db      42381 TCCCACACGGGACAATCTGGAAACGCTTCAAGAAGTCGTAGGGTG 42337
```

```
RESULT 11
AB097933/c
LOCUS       AB097933              125125 bp    DNA     linear   VRL 13-DEC-2002
DEFINITION  Human herpesvirus 3 ORF62 gene for IE62 transactivator, complete
            cds, strain: The Oka parental.
ACCESSION   AB097933
VERSION     AB097933.1  GI:26665422
KEYWORDS    .
SOURCE      Human herpesvirus 3
  ORGANISM  Human herpesvirus 3
            Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
            Alphaherpesvirinae; Varicellovirus.
REFERENCE   1
  AUTHORS   Gomi,Y., Sunamachi,H., Mori,Y., Nagaike,K., Takahashi,M. and
            Yamanishi,K.
  TITLE     Comparison of the complete DNA sequences of the Oka varicella
            vaccine and its parental virus
  JOURNAL   J. Virol. 76 (22), 11447-11459 (2002)
  MEDLINE   22276345
   PUBMED   12388706
REFERENCE   2  (bases 1 to 125125)
  AUTHORS   Gomi,Y., Sunamachi,H., Nagaike,K., Mori,Y., Takahashi,M. and
            Yamanishi,K.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-DEC-2002) Yasuyuki Gomi, Kanonji institute, the
            research foundation for microbial diseases of Osaka university,
            research division; 2-9-41 Yahata-cho, Kanonji-city, Kagawa
            768-0062, Japan (E-mail:ygomi@mail.biken.or.jp,
            Tel:81-875-25-4171(ex.620), Fax:81-875-25-4843)
FEATURES             Location/Qualifiers
     source          1. .125125
                     /organism="Human herpesvirus 3"
                     /mol_type="genomic DNA"
                     /strain="The Oka parental"
                     /db_xref="taxon:10335"
                     /note="synonym: varicella-zoster virus"
     gene            complement(105166. .109098)
                     /gene="ORF62"
     CDS             complement(105166. .109098)
                     /gene="ORF62"
                     /codon_start=1
                     /product="IE62 transactivator"
                     /protein_id="BAC44882.1"
                     /db_xref="GI:26665423"
                     /translation="MDTPPMQRSTPQRAGSPDTLELMDLLDAAAAAAEHRARVVTSSQ
                     PDDLLFGENGVMVGREHEIVSIPSVSGLQPEPRTEDVGEELTQDDYVCEDGQDLMGSP
                     VIPLAEVFHTRFSEAGAREPTGADRSLETVSLGTKLARSPKPPMNDGETGRGTTPPFP
                     QAPSPVSPASPVGDAAGNDQREDQRSIPRQTTRGNSPGLPSVVHRDRQTQSISGKKPG
                     DEQAGHAHASGDGVVLQKTQRPAQGKSPKKKTLKVKVPLPARKPGGPVPGPVEQLYHV
                     LSDSVPAKGAKADLPFETDDTRPRKHDARGITPRVPGRSSGGKPRAFLALPGRSHAPD
                     PIEDDSPVEKKPKSREFVSSSSSSSSWGSSSEDEDDEPRRVSVGSETTGSRSGREHAP
                     SPSNSDDSDSNDGGSTKQNIQPGYRSISGPDPRIRKTKRLAGEPGRQRQKSFSLPRSR
                     TPIIPPVSGPLMMPDGSPWPGSAPLPSNRVRFGPSGETREGHWEDEAVRAARARYEAS
                     TEPVPLYVPELGDPARQYRALINLIYCPDRDPIAWLQNPKLTGVNSALNQFYQKLLPP
                     GRAGTAVTGSVASPVPHVGEAMATGEALWALPHAAAAVAMSRRYDRAQKHFILQSLRR
                     AFASMAYPEATGSSPAARISRGHPSPTTPATQTPDPQPSAAARSLSVCPPDDRLRTPR
                     KRKSQPVESRSLLDKIRETPVADARVADDHVVSKAKRRVSEPVTITSGPVVDPPAVIT
                     MPLDGPAPNGGFRRIPRGALHTPVPSDQARKAYCTPETIARLVDDPLFPTAWRPALSF
                     DPGALAEIAARRPGGGDRRFGPPSGVEALRRRCAWMRQIPDPEDVRLLIIYDPLPGED
                     INGPLESTLATDPGPSWSPSRGGLSVVLAALSNRLCLPSTHAWAGNWTGPPDVSALNA
                     RGVLLLSTRDLAFAGAVEYLGSRLASARRRLLVLDAVALERWPRDGPALSQYHVYVRA
                     PARPDAQAVVRWPDSAVTEGLARAVFASSRTFGPASFARIETAFANLYPGEQPLCLCR
                     GGNVAYTVCTRAGPKTRVPLSPREYRQYVLPGFDGCKDLARQSRGLGLGAADFVDEAA
                     HSHRAANRWGLGAALRPVFLPEGRRPGAAGPEAGDVPTWARVFCRHALLEPDPAAEPL
                     VLPPVAGRSVALYASADEARNALPPIPRVMWPPGFGAAETVLEGSDGTRFVFGHHGGS
                     ERPAETQAGRQRRTADDREHALEPDDWEVGCEDAWDSEEGGGDDGDAPGSSFGVSIVS
                     VAPGVLRDRRVGLRPAVKVELLSSSSSSEDEDDVWGGRGGRSPPQSRG"
ORIGIN

  Query Match           99.8%;  Score 703.4;  DB 14;  Length 125125;
  Best Local Similarity   99.9%;  Pred. No. 4.2e-204;
  Matches  704;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 ATGACACAACCCGCATCGTCTCGTGTAGTCTTTGATCCCAGCAACCCCACCACATTTTCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      43017 ATGACACAACCCGCATCGTCTCGTGTAGTCTTTGATCCCAGCAACCCCACCACATTTTCG 42958

Qy         61 GTGGAAGCAATTGCGGCTTACACCCCCGTTGCTTTAATACGACTTTTAAACGCCAGTGGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42957 GTGGAAGCAATTGCGGCTTACACCCCCGTTGCTTTAATACGACTTTTAAACGCCAGTGGA 42898

Qy        121 CCTTTGCAACCTGGTCACCGTGTGGACATCGCTGATGCCAGAAGCATTTACACCGTGGGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42897 CCTTTGCAACCTGGTCACCGTGTGGACATCGCTGATGCCAGAAGCATTTACACCGTGGGA 42838

Qy        181 GCCGCGGCCAGTGCCGCGCGTGCACGCGCTAACCATAATGCAAATACGATACGCCGAACG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42837 GCCGCGGCCAGTGCCGCGCGTGCACGCGCTAACCATAATGCAAATACGATACGCCGAACG 42778

Qy        241 GCCATGTTTGCCGAGACTGACCCTATGACATGGTTAAGACCAACGGTTGGCTTAAAACGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42777 GCCATGTTTGCCGAGACTGACCCTATGACATGGTTAAGACCAACGGTTGGCTTAAAACGT 42718

Qy        301 ACGTTTAACCCGCGTATTATACGACCACAACCCCCAAATCCATCCATGAGTTTGGGAATC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42717 ACGTTTAACCCGCGTATTATACGACCACAACCCCCAAATCCATCCATGAGTTTGGGAATC 42658

Qy        361 TCGGGGCCTACTATATTGCCGCAAAAAACACAGAGCGCCGATCAGTCTGCTTTACAACAG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42657 TCGGGGCCTACTATATTGCCGCAAAAAACACAGAGCGCCGATCAGTCTGCTTTACAACAG 42598

Qy        421 CCCGCCGCGTTGGCGTTTTCGGGATCATCCCCGCAACACCCCCCACCTCAAACGACGTCG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42597 CCCGCCGCGTTGGCGTTTTCGGGATCATCCCCGCAACACCCCCCACCTCAAACGACGTCG 42538

Qy        481 GCATCCGTTGGACAACAGCAACACGTGGTGTCGGGGTCTTCTGGACAACAACCGCAACAG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42537 GCATCCGTTGGACAACAGCAACACGTGGTGTCGGGGTCTTCTGGACAACAACCGCAACAG 42478

Qy        541 GGAGCACAGTCAAGCACTGTCCAGCCAACAACCGGATCACCGCCCGCGGCCCAAGGCGTG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42477 GGAGCACAGTCAAGCACTGTCCAGCCAACAACCGGATCACCGCCCGCGGCCCAAGGCGTG 42418

Qy        601 CCACAGTCTACCCCGCCCCCAACCCAAAATACCCCCCAGGGGGGTAAGGGACAGACCTTG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42417 CCACAGTCTACCCCGCCCCCAACCCAAAATACCCCCCAGGGGGGTAAGGGACAGACCTTG 42358

Qy        661 TCACACACGGGACAATCTGGAAACGCTTCAAGAAGTCGTAGGGTG 705
              || |||||||||||||||||||||||||||||||||||||||||||
Db      42357 TCCCACACGGGACAATCTGGAAACGCTTCAAGAAGTCGTAGGGTG 42313
```

```
RESULT 12
AR431492/c
LOCUS       AR431492              125157 bp    DNA     linear   PAT 18-DEC-2003
DEFINITION  Sequence 2 from patent US 6653069.
ACCESSION   AR431492
VERSION     AR431492.1  GI:40193596
```

```
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 125157)
  AUTHORS   Gomi,Y., Sunamachi,H., Takahashi,M. and Yamanishi,K.
  TITLE     Method for quality control of an attenuated vericella live vaccine
  JOURNAL   Patent: US 6653069-A 2 25-NOV-2003;
FEATURES             Location/Qualifiers
     source          1. .125157
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match             99.8%;  Score 703.4;  DB 6;  Length 125157;
  Best Local Similarity   99.9%;  Pred. No. 4.2e-204;
  Matches  704;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 ATGACACAACCCGCATCGTCTCGTGTAGTCTTTGATCCCAGCAACCCCACCACATTTTCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      43113 ATGACACAACCCGCATCGTCTCGTGTAGTCTTTGATCCCAGCAACCCCACCACATTTTCG 43054

Qy         61 GTGGAAGCAATTGCGGCTTACACCCCCGTTGCTTTAATACGACTTTTAAACGCCAGTGGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      43053 GTGGAAGCAATTGCGGCTTACACCCCCGTTGCTTTAATACGACTTTTAAACGCCAGTGGA 42994

Qy        121 CCTTTGCAACCTGGTCACCGTGTGGACATCGCTGATGCCAGAAGCATTTACACCGTGGGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42993 CCTTTGCAACCTGGTCACCGTGTGGACATCGCTGATGCCAGAAGCATTTACACCGTGGGA 42934

Qy        181 GCCGCGGCCAGTGCCGCGCGTGCACGCGCTAACCATAATGCAAATACGATACGCCGAACG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42933 GCCGCGGCCAGTGCCGCGCGTGCACGCGCTAACCATAATGCAAATACGATACGCCGAACG 42874

Qy        241 GCCATGTTTGCCGAGACTGACCCTATGACATGGTTAAGACCAACGGTTGGCTTAAAACGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42873 GCCATGTTTGCCGAGACTGACCCTATGACATGGTTAAGACCAACGGTTGGCTTAAAACGT 42814

Qy        301 ACGTTTAACCCGCGTATTATACGACCACAACCCCCAAATCCATCCATGAGTTTGGGAATC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42813 ACGTTTAACCCGCGTATTATACGACCACAACCCCCAAATCCATCCATGAGTTTGGGAATC 42754

Qy        361 TCGGGGCCTACTATATTGCCGCAAAAAACACAGAGCGCCGATCAGTCTGCTTTACAACAG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42753 TCGGGGCCTACTATATTGCCGCAAAAAACACAGAGCGCCGATCAGTCTGCTTTACAACAG 42694

Qy        421 CCCGCCGCGTTGGCGTTTTCGGGATCATCCCCGCAACACCCCCCACCTCAAACGACGTCG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42693 CCCGCCGCGTTGGCGTTTTCGGGATCATCCCCGCAACACCCCCCACCTCAAACGACGTCG 42634

Qy        481 GCATCCGTTGGACAACAGCAACACGTGGTGTCGGGGTCTTCTGGACAACAACCGCAACAG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42633 GCATCCGTTGGACAACAGCAACACGTGGTGTCGGGGTCTTCTGGACAACAACCGCAACAG 42574

Qy        541 GGAGCACAGTCAAGCACTGTCCAGCCAACAACCGGATCACCGCCCGCGGCCCAAGGCGTG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42573 GGAGCACAGTCAAGCACTGTCCAGCCAACAACCGGATCACCGCCCGCGGCCCAAGGCGTG 42514

Qy        601 CCACAGTCTACCCCGCCCCCAACCCAAAATACCCCCCAGGGGGGTAAGGGACAGACCTTG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42513 CCACAGTCTACCCCGCCCCCAACCCAAAATACCCCCCAGGGGGGTAAGGGACAGACCTTG 42454

Qy        661 TCACACACGGGACAATCTGGAAACGCTTCAAGAAGTCGTAGGGTG 705
              || |||||||||||||||||||||||||||||||||||||||||||
Db      42453 TCCCACACGGGACAATCTGGAAACGCTTCAAGAAGTCGTAGGGTG 42409


RESULT 13
BD097171/c
LOCUS       BD097171              125157 bp    DNA     linear   PAT 27-AUG-2002
DEFINITION  Method for the quality control of attetunated varicella live
            vaccine.
ACCESSION   BD097171
VERSION     BD097171.1  GI:22642745
KEYWORDS    WO 0156600-A/2.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 125157)
  AUTHORS   Gomi,Y., Sunamachi,H., Takahashi,M. and Yamanishi,K.
  TITLE     Method for the quality control of attetunated varicella live
            vaccine
  JOURNAL   Patent: WO 0156600-A 2 09-AUG-2001;
            THE RESEARCH FOUNDATION FOR MICROBIAL DISEASES OF OSAKA UNIVERSITY,
            YASUYUKI GOMI,HIROKI SUNAMACHI,MICHIAKI TAKAHASHI,KOICHI YAMANISHI
COMMENT     OS   Attenuated varicella virus Oka strain
            PN   WO 0156600-A/2
            PD   09-AUG-2001
            PF   31-JAN-2001 WO 2001JP000678
            PR   31-JAN-2000 JP 00P    62734
            PI   YASUYUKI GOMI,HIROKI SUNAMACHI,MICHIAKI TAKAHASHI,KOICHI  PI
            YAMANISHI
            PC   A61K39/25,A61P31/22,C12N15/38,C12N7/00,G01N33/15 CC   Method
            for the quality control of attetunated varicella live  CC
             vaccine
            FH   Key             Location/Qualifiers
            FT   source          1. .125157
            FT                   /organism='Attenuated varicella virus Oka  FT
                                 strain'.
FEATURES             Location/Qualifiers
     source          1. .125157
                     /organism="unidentified"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32644"
ORIGIN

  Query Match             99.8%;  Score 703.4;  DB 6;  Length 125157;
  Best Local Similarity   99.9%;  Pred. No. 4.2e-204;
  Matches  704;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 ATGACACAACCCGCATCGTCTCGTGTAGTCTTTGATCCCAGCAACCCCACCACATTTTCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      43113 ATGACACAACCCGCATCGTCTCGTGTAGTCTTTGATCCCAGCAACCCCACCACATTTTCG 43054

Qy         61 GTGGAAGCAATTGCGGCTTACACCCCCGTTGCTTTAATACGACTTTTAAACGCCAGTGGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      43053 GTGGAAGCAATTGCGGCTTACACCCCCGTTGCTTTAATACGACTTTTAAACGCCAGTGGA 42994

Qy        121 CCTTTGCAACCTGGTCACCGTGTGGACATCGCTGATGCCAGAAGCATTTACACCGTGGGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42993 CCTTTGCAACCTGGTCACCGTGTGGACATCGCTGATGCCAGAAGCATTTACACCGTGGGA 42934

Qy        181 GCCGCGGCCAGTGCCGCGCGTGCACGCGCTAACCATAATGCAAATACGATACGCCGAACG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42933 GCCGCGGCCAGTGCCGCGCGTGCACGCGCTAACCATAATGCAAATACGATACGCCGAACG 42874

Qy        241 GCCATGTTTGCCGAGACTGACCCTATGACATGGTTAAGACCAACGGTTGGCTTAAAACGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42873 GCCATGTTTGCCGAGACTGACCCTATGACATGGTTAAGACCAACGGTTGGCTTAAAACGT 42814

Qy        301 ACGTTTAACCCGCGTATTATACGACCACAACCCCCAAATCCATCCATGAGTTTGGGAATC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42813 ACGTTTAACCCGCGTATTATACGACCACAACCCCCAAATCCATCCATGAGTTTGGGAATC 42754

Qy        361 TCGGGGCCTACTATATTGCCGCAAAAAACACAGAGCGCCGATCAGTCTGCTTTACAACAG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42753 TCGGGGCCTACTATATTGCCGCAAAAAACACAGAGCGCCGATCAGTCTGCTTTACAACAG 42694

Qy        421 CCCGCCGCGTTGGCGTTTTCGGGATCATCCCCGCAACACCCCCCACCTCAAACGACGTCG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42693 CCCGCCGCGTTGGCGTTTTCGGGATCATCCCCGCAACACCCCCCACCTCAAACGACGTCG 42634

Qy        481 GCATCCGTTGGACAACAGCAACACGTGGTGTCGGGGTCTTCTGGACAACAACCGCAACAG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42633 GCATCCGTTGGACAACAGCAACACGTGGTGTCGGGGTCTTCTGGACAACAACCGCAACAG 42574
```

```
Qy        541 GGAGCACAGTCAAGCACTGTCCAGCCAACAACCGGATCACCGCCCGCGGCCCAAGGCGTG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42573 GGAGCACAGTCAAGCACTGTCCAGCCAACAACCGGATCACCGCCCGCGGCCCAAGGCGTG 42514

Qy        601 CCACAGTCTACCCCGCCCCCAACCCAAAATACCCCCCAGGGGGGTAAGGGACAGACCTTG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42513 CCACAGTCTACCCCGCCCCCAACCCAAAATACCCCCCAGGGGGGTAAGGGACAGACCTTG 42454

Qy        661 TCACACACGGGACAATCTGGAAACGCTTCAAGAAGTCGTAGGGTG 705
              || ||||||||||||||||||||||||||||||||||||||||||
Db      42453 TCCCACACGGGACAATCTGGAAACGCTTCAAGAAGTCGTAGGGTG 42409


RESULT 14
AF275348/c
LOCUS       AF275348               124138 bp    DNA     linear   VRL 07-MAR-2002
DEFINITION  Cercopithecine herpesvirus 7, complete genome.
ACCESSION   AF275348
VERSION     AF275348.2  GI:13186283
KEYWORDS    .
SOURCE      Cercopithecine herpesvirus 7
  ORGANISM  Cercopithecine herpesvirus 7
            Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
            Alphaherpesvirinae; Varicellovirus.
REFERENCE   1  (bases 1 to 124138)
  AUTHORS   Gray,W.L., Starnes,B., White,M.W. and Mahalingam,R.
  TITLE     The DNA sequence of the simian varicella virus genome
  JOURNAL   Virology 284 (1), 123-130 (2001)
  MEDLINE   21251382
   PUBMED   11352673
REFERENCE   2  (bases 1 to 124138)
  AUTHORS   Gray,W.L., Starnes,H.B., White,M.W., Ashburn,C.V. and Mahalingam,R.
  TITLE     Complete Sequence of the Simian Varicella Virus Genome
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 124138)
  AUTHORS   Gray,W.L., Starnes,H.B., White,M.W., Ashburn,C.V. and Mahalingam,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-JUN-2000) Microbiology/Immunology, University of
            Arkansas for Medical Sciences, 4301 W. Markham St., Little Rock, AR
            72205, USA
REFERENCE   4  (bases 1 to 124138)
  AUTHORS   Gray,W.L., Starnes,H.B., White,M.W., Ashburn,C.V. and Mahalingam,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-MAR-2001) Microbiology/Immunology, University of
            Arkansas for Medical Sciences, 4301 W. Markham St., Little Rock, AR
            72205, USA
  REMARK    Sequence update by submitter
COMMENT     On Mar 2, 2001 this sequence version replaced gi:11036551.
FEATURES             Location/Qualifiers
     source          1. .124138
                     /organism="Cercopithecine herpesvirus 7"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:35245"
     misc_feature    8. .9
                     /note="TRL/UL boundary"
     CDS             complement(926. .1807)
                     /note="ORFA"
                     /codon_start=1
                     /product="unknown"
                     /protein_id="AAG27248.1"
                     /db_xref="GI:11036619"
                     /translation="METKRFRRKQLSDSTSPLHNTQRAIRSAGLNVSGTEISKFIMFM
                     KTFMKKMRAVYLEPQRKHNVINGGNEMWKCGGCLPTTDDDVSWDDLVLYGHNVYEAFQ
                     KIQSNVNISCALRERLLRGETLCEALTSADELLTWVKLMAVKKLPIWTWNPIVATSSV
                     LLDNLKLKLEPMVKCILTQQRAETTTLVELLERQSFNDITCITTLMFVTVARIVTAVQ
                     RGDEHIAYENVSESCNVLEDYTPGACEAGVLGALMHHGRYCNKVECKIVCWPEYVHHT
                     EKFGKFFRCSEFNIYCK"
     CDS             2327. .2671
                     /note="ORFB"
                     /codon_start=1
                     /product="unknown"
                     /protein_id="AAG27181.1"
                     /db_xref="GI:11036552"
                     /translation="MSRSPTVERFILSSTELTLSSFSPITYNAEYMNVYLPTYEEAIQ
                     DLPPAYRSRETLSTLTTDTGCMDCICTGLYKIHERLTSCVRYCVPAMFVLFAVLTLTA
                     VILVILAAIPDV"
     CDS             complement(2731. .3036)
                     /note="ORF1"
                     /codon_start=1
                     /product="unknown"
                     /protein_id="AAG27182.1"
                     /db_xref="GI:11036553"
                     /translation="MSWVSRGSKYAEESDGDDSVFTYSCDMNNQRSEYINRTVFPTET
                     VIPLFNASHSNIKPEWRKTKGGNKQKLWNVVFICIQVCTLMAVMYLIYKSVSPSNRR"
     CDS             complement(3301. .3852)
                     /note="ORF3"
                     /codon_start=1
                     /product="unknown"
                     /protein_id="AAG27203.1"
                     /db_xref="GI:11036574"
                     /translation="MACNQSVNKETLMIDLKPDALAPFTQAVAPLALETTWISATSSH
                     APTPSPLYGTKRLCALRATCTKPGDVHAYLIGLSHKDKPSEFPLFVDIQNLCKILQCQ
                     RLLKEMANYSALCDAPFSAATLQIMLDSGSLGVHLSGVGYHCHCKTPFSVDCWAGAAE
                     AYDHIVCTLRVRNVMGISPENRD"
     CDS             complement(4121. .5533)
                     /note="ORF4"
                     /codon_start=1
                     /product="transactivator"
                     /protein_id="AAG27214.1"
                     /db_xref="GI:11036585"
                     /translation="METASTTSCDSDIMACVCDDSMHLSPEDDPIDDFQLEVQNWAEE
                     SVPNSVQIDDDEEAGEILTVTEETPNPADVCEDADRVYTQPNFESKALSGYHKDDEKC
                     LVTFSDNIPRRTNPIHVHYTNGERKSPKWYNQSYRKYQRRESWRYKHGRYSGRPDRSP
                     SKHFRRTTDAEYVRYNERRSRRDYIRDISTDKLTEQARSTLKIAGLSISTNSVLEFTN
                     VMKALIRRMTTFYDEDGNIEKHESWKDGGCYPTTEENISWAELMLYGHKLYQTFQTYR
                     VSIKLGRALRERIIRGESLLEALESADELITWLKLMAAKNLPIWTWNPIVATSKSLVE
                     NLKLKLGPVMRCLLLRRDGGEELSTTFPELLHRQRFSDITCMPTFMFVVISRVVTAVL
                     RGDKCIAYEDISSSAQVLEEYTPGSCKAGVLESIITHGRYCGIPECRITCWPEFVPET
                     HHHGKFFKCSYFENTVNAPN"
     CDS             complement(5643. .6656)
                     /note="ORF5"
                     /codon_start=1
                     /product="unknown"
                     /protein_id="AAG27224.1"
                     /db_xref="GI:11036595"
                     /translation="MQALGIKTEHFIIVCLLMTYLLFVLWYTSRIQFRNVCVYATAPV
                     QGGKVIWNLYNSSLIYVTVTNHSLFLDGLSGYDDSCRKDLINGDSILPTVFSSPLHEK
                     IRVVLGTRNCRAYMWCVHLKIIIINLFVYAVYLQCRRVRRMFGPFRKACEFISPMSYS
                     LQYTTSVISTTFLKFPYNKLTKLLHEVESQRNAISSMFMYDPISFLCLNKLIGLLIAF
                     EIFVHIISGCVLLTLGIVYTPCSLLYPLYVRIVAWMFTGLIFGVEMVACVRSKSNNEN
                     NENNGSTVNPKGLHGVCTACCATVISGLAVKCIYVAVFAIAVIVFMHYEQRVQVSLFW
                     HIR"
     CDS             complement(6661. .9906)
                     /note="ORF6"
                     /codon_start=1
                     /product="DNA helicase-primase complex component"
                     /protein_id="AAG27235.1"
                     /db_xref="GI:11036606"
                     /translation="MEALEKSQTIRILFATKSCAVSYCLQLLTGQISEDPVYVLSYTW
                     TADLNDVFYEDGWADKEVPTKRIRLDGPVKTISSYNYLHGKRVRVIFYLIGGVITGAI
                     KARIWPVFVCEFSKEGDLTNLCKSLWFGEPLQTRVLSETLHELDTFAIHADIIQCLLT
                     THRASPKTTHCKVDANYKDNGLIQLNKKMERDHLISLYAHYERNVLAAYRRIYWGYSC
                     TPFWYVARFGPYEKTLVLAMRYYFLQTETSYMTVTDKVYDLQAIKDLLETYNVPERAN
                     LSGYSISDLITFEKLSAFCCSSMYTRGIVAINAPLYISSRISIDVSEIESVITYITND
                     RKGLRVSDREFVDYIYLAYFECFNRKQIKEHLNAVICNSVEQVPIIANMSRLGLQAVD
                     NFFTHVRSRVNINDYISKNVICDFRTLTNAMAVKYIKEKTYAIVNKSVCSRLDVVGIC
                     DSAISLTHRIEDIEKVFGKLGWLSTKSFINNVCFSSNDNESDVSVEAIKAANQANSIN
                     SHKPAWEYNHSITRTLLFFCSTNIQGNNSLTATYKSPFQALLGCTKISGPIPVYRISL
                     PQGRQSFALIVYDFWKCVTDIYTQNEMLSVMESITGKEKLPLCFRDMKVTNELLATVK
                     CSGIASSIPDEQMYINRNELFNSTLAVSNIILDVDFHVRKAIPKGVLHAGMRAFRHGA
                     LTAVQLLFCNSAVDINSYPCYFYKSSCNITESMCDVDENCDIVENYNVDEYNLYETVE
                     MNDVDDEMWLCDIVNNIQQYTKSEESETIPDIYNTNTCSCNKNLGFRVCIPVPPPYVL
                     TGPSTLKAIAKIVQQAVLLENDFVKHMGKYLKDFSFIDTGVYSHGHSLRLPFFSKVTH
                     NTTFYGRLLPFYVIPDVCSDILSFISAHSNPENFHFHSKPSFNYDTIIVLHNITGEYA
                     DFFERKVAYNKEVFGTPKVSLEKTLQSHGLETVDAHTIEAFIDSIVIESIIQYIHLHF
                     PGRREREYSMTSSKCIVVKRDWVLLQLTHCLKTSGRQGFTCLRYPHRGGRTAPRNFISL
```

```
                RVDHHNRLSVSLSQQCFATKCDSNRMHTIFTLQLPNGTRGDVNHITETIET"
  CDS           9917. .10612
                /note="ORF7"
                /codon_start=1
                /product="unknown"
                /protein_id="AAG27244.1"
                /db_xref="GI:11036615"
                /translation="MHRICAGLCGYTRVPLEDTSYEEVRGKPGPSGPALLRLQEALAM
                VNMLLPAPVTIEDVVTSADNTRRLVRAQALARTYAACSRNLECLRQHKFSEENPSLNS
                VIKSHIQDSQRLADTCLAAITNLYLSVGSVDTTTDALVDQTIRMVAENEVVMSDVILM
                EKTLGVDNQCGVNAMLTPAPPGNIAPCSEDTERHNITKPFTGKASRNKRVSEESSQYH
                SNVPSRKPEMEHV"
  repeat_region 10761. .10872
                /note="reiteration 1"
  CDS           complement(10881. .12068)
                /note="ORF8"
                /codon_start=1
                /product="dUTPase"
                /protein_id="AAG27245.1"
                /db_xref="GI:11036616"
                /translation="MNNSKEQFEAMIAAAENRERTGILERRRQRNEYYRDEILLDVGM
                RNNSDFNCALDGDCGWNQPVVLWNRRKITLPSDEAMECGDFYIQRLDLGVKASTPNGY
                ALLLLRCIHSNPDKNTDVFIANGVIDSGYRGRICAVLYYKKNVTIIRPGDLMVYLLPV
                KLATSYVLNVVYDNSLDPVLAEHMCVGKEKIAMKLFNDVNNAIKTGEQCIKSHCNYTE
                THVPIRLIFSNTLFKSKRVEDAGYDVCAPFEITVPARTTIRVTVPFIQRLNVNNMDAY
                IFGRSSKNVIGVIVLPTVWIWGSLCEFYIFNATSSNVIVKPGEKIAQVVLLEHRNQSV
                HIHKDIINNFEPFPSIIFGNHTIEKLNSPDVYPKWHFTTMFNYIAPPSDRGNKGFGST
                DVE"
  CDS           12037. .12300
                /note="ORF9A"
                /codon_start=1
                /product="membrane protein"
                /protein_id="AAG27247.1"
                /db_xref="GI:11036618"
                /translation="MASNCSLELFMLLLLSVFTSAHTTHKLPEHRNFWHASCSARGVY
                IDGSAIATLFFYASLLGVCVALTALAYNACFRLFTRSVLRSTW"
  CDS           12405. .13310
                /note="ORF9"
                /codon_start=1
                /product="unknown"
                /protein_id="AAG27246.1"
                /db_xref="GI:11036617"
                /translation="MASPDGDRIRRSNAVRRKHMDAQYVNVPYVGKPRNPVPRRSVVV
                GFPDDSDDSLGYIATVETESPPLFYGDTSGTVDKRHTRMISNNSSDSDDDFEDIDEVI
                AAFKEARLKHEHVEDAVYENVQHGNAQPRMIQKHNRDVDSPKRAPPGAGAIASGRPLS
                FSTAPKNTTSAWCGPTPAYNKRVFCEAIRRIAAAQAQRAAEAAWSSNPPRNNAELDRL
                LAGTVVRITVHEGLNLIQVANDLELGGKPDRKQSTARRKKEIVKDTDDEPVYAQVKIN
                SRGTQPVSGTSRARTRSVSKPDGRK"
  CDS           13537. .14757
                /note="tegument protein; ORF10"
                /codon_start=1
                /product="transactivator"
                /protein_id="AAG27183.1"
                /db_xref="GI:11036554"
                /translation="MEHFNGAFDVEQWHNRHLEQAFLNDFDESLLGDVQEHNAGSLYS
                SVQVGTPPPRVASPAVLYQQLINDLHFSEGPRIIALLETCNEDIFSCFPLNEDLYTDT
                IVFSTNPSDIIESLETSETPEPINLITKGCDLPSLPSTYTELPMYVENVQNTFFTELR
                VREKAYLRLLILYCKSLLRHLQGISRRSSVGARSQDCEHKYITQFKQSIRLRYYREVA
                KLARLLYLHLYITVTRQFSWRLFAQQSKYPDVFAAMKFMWLENRQFTCAFHPILCNHG
                IVLLEGKPLSSQRLREVNYRRRQLGLPLVRCALIEEKGVPLVQRPEFSAHMPRSLGFL
                THHIKAKLDVYSSKHSQEPGNVRVDHPYARNVNNINYGSSVEAMIVSPPSPSDILPGD
                PPQKLKIDVVSDVP"
  CDS           15076. .17004
                /note="ORF11"
                /codon_start=1
                /product="unknown"
                /protein_id="AAG27184.1"
                /db_xref="GI:11036555"
                /translation="MDNSNTISIPSNSDNEEGELIDSDNDDGIFIYMDAYQPPNENFS

Query Match             25.7%;  Score 181.2;  DB 14;  Length 124138;
Best Local Similarity   71.8%;  Pred. No. 1e-43;
Matches  237;  Conservative    0;  Mismatches   93;  Indels    0;  Gaps    0;

Qy          1 ATGACACAACCCGCATCGTCTCGTGTAGTCTTTGATCCCAGCAACCCCACCACATTTTCG 60
              ||||||||| |  |||||||||| |      |||||||        ||   ||| |||||
Db      43870 ATGACACAATCTACATCGTCTCGAGCGCATTTTGATCCATCTGTACCAGGCACCTTTTCT 43811

Qy         61 GTGGAAGCAATTGCGGCTTACACCCCCGTTGCTTTAATACGACTTTTAAACGCCAGTGGA 120
              ||||| |||||||| ||  |||  || ||||| ||| ||| |||  ||||||| ||||||
Db      43810 GTGGATGCAATTGCAGCAAACAGTCCTGTTGCATTATTACAACTAGTAAACGCTAGTGGA 43751

Qy        121 CCTTTGCAACCTGGTCACCGTGTGGACATCGCTGATGCCAGAAGCATTTACACCGTGGGA 180
              ||| |     |||||   ||| | || || |   ||    | || || ||||||||| ||
Db      43750 CCTGTTATTTCTGGTGTTCGTTTTGATATTGGACATATTCGGAGTATATACACCGTAGGT 43691

Qy        181 GCCGCGGCCAGTGCCGCGCGTGCACGCGCTAACCATAATGCAAATACGATACGCCGAACG 240
              || || |||  ||||||  |  | || |  |||||||||| |||||| || ||| ||||
Db      43690 GCTGCTGCCGCTGCCGCTAGATCTCGTGTAAACCATAATGTAAATACTATTCGCAGAACT 43631

Qy        241 GCCATGTTTGCCGAGACTGACCCTATGACATGGTTAAGACCAACGGTTGGCTTAAAACGT 300
              || |||||||| || || ||| |||||||||| |  |||||||||||| || ||| |
Db      43630 GCGATGTTTGCAGAAACGGACGCTATGACATGGCTTCGACCAACGGTTGGTTTGAAAAGA 43571

Qy        301 ACGTTTAACCCGCGTATTATACGACCACAA 330
              |||||||| ||||||||| | || || |||
Db      43570 ACGTTTAATCCGCGTATTGTTCGTCCGCAA 43541


RESULT 15
AF282130
LOCUS       AF282130               151102 bp    DNA     linear   VRL 14-MAY-2001
DEFINITION  Meleagrid herpesvirus 1 unique long region, IRL region, IRS region,
            unique short region, complete sequences and IRS partial sequence.
ACCESSION   AF282130
VERSION     AF282130.1  GI:11095830
KEYWORDS    .
SOURCE      Meleagrid herpesvirus 1 (herpesvirus of turkeys)
  ORGANISM  Meleagrid herpesvirus 1
            Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
            Alphaherpesvirinae; Mardivirus.
REFERENCE   1  (bases 1 to 151102)
  AUTHORS   Kingham,B.F., Zelnik,V., Kopacek,J., Majerciak,V., Ney,E. and
            Schmidt,C.J.
  TITLE     The genome of herpesvirus of turkeys: comparative analysis with
            Marek's disease viruses
  JOURNAL   J. Gen. Virol. 82 (Pt 5), 1123-1135 (2001)
  MEDLINE   21195611
   PUBMED   11297687
REFERENCE   2  (bases 1 to 151102)
  AUTHORS   Schmidt,C.J., Zelnik,V., Kingham,B., Kopacek,J., Majerciak,V.,
            Ney,E. and Chen,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JUN-2000) Animal and Food Sciences, University of
            Delaware, Townsend Hall, 531 S. College Ave., Newark, DE 19717, USA
FEATURES             Location/Qualifiers
     source          1. .151102
                     /organism="Meleagrid herpesvirus 1"
                     /mol_type="genomic DNA"
                     /strain="FC126"
                     /db_xref="taxon:37108"
     repeat_region   <1. .97
                     /note="terminal repeat long; TRL"
     misc_feature    98. .110792
                     /note="long unique region; UL"
     gene            527. .2788
                     /gene="LORF1"
     CDS             join(527. .622,698. .2788)
                     /gene="LORF1"
                     /codon_start=1
                     /product="phospholipase-like protein"
                     /protein_id="AAG30038.1"
                     /db_xref="GI:11095831"
                     /translation="MPAKGEVEHHAGYGVALAILALLLIHATALIMIFSDVNLRTGQH
                     GDDVHLRYPLTFPDGATPPRAGDHGHSELWTDIKNNTDSTYSSQPPTMLFPKPTTVSHY
                     PEHNSTTIIIPASEMPAMQDIVSIVVSKRHSGSRDAWHVSEYGPNIIRMPHTHKDIPK
```

PFNSILSNQFALLTMWDNLVGMFKTFWGNAMGFHQMIREGLVDLRLHTDVIFLVHGWH
GRTNDTEFFLSANRLLTRMIPSAAVVYVSWDPQGAMEPAADAILLARRINVTQFLYAV
PSNMGIQCIGHSLGSYVCGSICRQYYYTGKGRCKRILGINAYEILSTDSSLISRIQSD
TIRSDAEYVSIFSMGERYIRTVYKDADEYIVAKGISRGVSTCVNPFAWKFSLCTRGVK
GREYCENFGANNVTFTSGLREDGSEICERLIPVMAVLQTLDVSSTYALLQLDPKRPSG
MTAHISSVWNVYVLEKDYRYATYAKNEPVWYSTVLSVGPNGFTPAAVFAVLSRSDTPV
IIRDAIQQRYSLWGNISMHSAFIEDSPTHFPSIGIRIQSPILGAFAWKGRVYSGPRIQ
IPVPEQAVVDYNCLPKDADHLCVSTKNVSSITIWRKLLLANYSFSILSIDDCLPYIPH
MAVIWKHPAIEVATWRNITVTFHPSVQLMAMRFENVLTPRNYTLMTFHDVCRADEINS
GIFFEYDWLGANMTIAVSKAGSYVLRWFFEMGIYEIPIAVYENERE"

gene 3543. .4778
/gene="LORF2"

CDS 3543. .4778
/gene="LORF2"
/codon_start=1
/product="LORF2"
/protein_id="AAG30039.1"
/db_xref="GI:11095832"
/translation="MPSAGFKQPAAVAIVDGLFAACRGLAVECMCNGYATVELMEKLA
TSWLTSVSAMVTWIPMFAREAEPPNQPEKNNRDARAYIIKLSGKINVNHATWPFWKLM
GKCLDWCCSFHAPNDTSCTHGSPRIGIRQRGNNPTFTPALDLYSSIMTWTPMPCFAET
REYQDSGDDDGPSTSRGSGVARVKPTVIRHRVDKTSPSRYENHNARTFAMANQDKRSI
PSWVHETDAAAQRPGQASPSRDNGRGSVATNSRSSSSPTISNVRTLDAAAATNLGWEA
PSPPRTPTLDPEPAIAGPSRAVNRSPSPRPLSPPPYQPLSWDMNAQTWDSSAGSDDDG
SAIDSTGGLHHTMRIVPHSTSGDRSGRDRNRSRSRSRSRSRSGSTERARRYRRRTGRG
VPRDQSESTVIVSSPEASD"

gene 5078. .5632
/gene="UL1"

CDS 5078. .5632
/gene="UL1"
/codon_start=1
/product="UL1 glycoprotein gL precursor"
/protein_id="AAG30040.1"
/db_xref="GI:11095833"
/translation="MFTAFFLQSAEAFMNWSDRDLTYGFVRAPNVTTIVHLECIPTSK
LTSMRYAEPSSDEVPSGIIIKTNCSLPEFILWYERVGVAAWVNPIIGTSLLLEDVLRS
RLDDSVKAGIGTLLSKIAYLIPTSHLRNRGAGCINLYASHDGTCYGSVHFDRFERSAD
DDNRGSFCRNKTPRLRNGGRPRET"

gene 5505. .6452
/gene="UL2"

CDS 5505. .6452
/gene="UL2"
/codon_start=1
/product="UL2 uracil-DNA glycosylase"
/protein_id="AAG30041.1"
/db_xref="GI:11095834"
/translation="MARVMARCTSIDSSEAPMTTIGDLSAETKHPDLEMVEGLGRPRK
RRRPEGLPPGFIPARTKSDKSCDGGEESSHVCNVETESKKAWEHFSSVYNIDCRWKEI
LGPELCSPTGTKILAEYERRLRYEEVYPSKSDVFAWTRYCAPEHVKVVIVGQDPYANE
GQAHGLAFSVPPGSPAPPSLKNILAAVRACFPSIEIGSSGCLEDWARRGVLLLNSVLT
VKRGDPGSHRAVVWQILVRSVLRKLSQSNAGLVFMLWGAQAQTMYYQNDNDGRHLVLK
YSHPSPLSKKPFVHCTHFRDANEFLCKMGKEAIDWSLTA"

gene 6527. .7168
/gene="UL3"

CDS 6527. .7168
/gene="UL3"
/codon_start=1
/product="UL3"
/protein_id="AAG30042.1"
/db_xref="GI:11095835"
/translation="MACNAPSVMAVLADWGWSCSTSGAAGYEQASLGRGCAREGIDYQ
GCGDPSSAGHELPATSADASDGVVANPEYTHFDTLFMVSSIDELGRRQLTDTIRRDLR
HSLAKFTIACTKTSSFSSSNATRKTRGRISHRNFQSNKSLQMFILGRRAHAKHIRAQL
QSVIQARKPRKYYTRAVDGSTRPVVPVFVYEFTAIDTVSLHRDNVIEIDAPNP"

gene complement(7445. .8275)
/gene="UL4"

CDS complement(7445. .8275)
/gene="UL4"
/codon_start=1
/product="UL4"
/protein_id="AAG30043.1"
/db_xref="GI:11095836"
/translation="MASTFIAYTLENARMSPICNTPPYEQIICSCETGTRSIAVGRMT
RCDHIPPSNFIIQRGPVGTLIVVDSGSIIYSYLLRVEEYTAEHHSASLSSLPEALCVI
PFATCTVMGTDAYVYNNSGGILTILGMGNSIYVTITIYGQSGILAKPCDNFQMSQQMG
PRRTDDEIVQKNDGSQNIDDVETITLESEYPGWRTTFTIAGHTNNIPGTAEVSDARED
LLAEAIREARLYISNPSCISDAGTEYEMHDGSLTNLHVPESIAKCYEPCIDQSPVYRK
"

gene complement(8349. .10925)
/gene="UL5"

CDS complement(8349. .10925)
/gene="UL5"
/codon_start=1
/product="UL5 DNA helicase/primase associated protein"
/protein_id="AAG30044.1"
/db_xref="GI:11095837"
/translation="MARSSTDRFDESMYLNFTAMHGIHSIISRVRVLANATLPYELIP
PLSYFIEASNVEDPSTLSLRDLPFSAYLISGNAGSGKSTCIQTLSEIMDCIITGTTKV
ASQNIYGKLSNSYTSSHINTIFQEFGFKGHHVQATLGKWHYACTSNPPTIKELQKRDL
VYYWEVLSDITKCTLKLLETDAGPRKFDMVYMLEEVLGRDRGSLAWTAFGVNGSLPSF
TRSNIIVIDEAGLIGKHLLTAIVYCWWLINATYRTPQYIKGRRPVLVCVGSPTQTSSL
ETTFEHSKLRCNVRTSENILTYIICNPTLRSYLDLSKNWSIPINNKRCTEPEFGDLLK
VLEYGLPITEEHARIVDSFVVPESYINNPANLPGWTRLYSSHKEVSAYMSRLHDYLKT
SGDNMFVVFTLPAYTFINLEAFEKYCTAANQTSITIEKWLVANASRLSNWSQSRDQDS
TRSKCEIRSQLGLAISCSDITYVLNSQVAVTTRLRKWVFGFCGTFESFLSVLKDDSFI
KTHGEGSIEYAYRFLSQLLFNGMINFYNYLQQRGLQENVVKVAYERLATLTTSILFPG
GQTQTHDITDVHESAILTGINCREYSQSASAPFESHWGEADDVIFSALDGQMIDLLYC
NYEFTRLESSSEIYAQFSMLKNIFTGRYAIMTDLFGDKFSKHRFDSYVDNVSTKGCQI
FINNMRGGMSSLALQTDSYTLMGYTYARVNAFADEPIRRKIDTCVAEVLCELNMPTVV
LKDQHGFMAALNSNITDFVEMVDENELKMAVTADYGISSKLAMTIARSQGLSLERAAI
CFPQSGLKLSSVYVAMSRVTSSKFLRMNINPLRETHTRDDNNISEHLLAALRDPNVHI
VY"

gene 10962. .13151
/gene="UL6"

CDS 10962. .13151
/gene="UL6"
/codon_start=1
/product="UL6 minor capsid protein"
/protein_id="AAG30045.1"
/db_xref="GI:11095838"
/translation="MMRRHTAHGRCDILRAIGDGNMDYDPHRDKIRRPESESVTKPAV
LSNAPSEPCVENEWILVHPTKQTCLFKEILSGQWGYTEGQGVYNSIRSAEAVSRQLQS
TIITLTLEAARYDELHRDWMKHIHARGMSAKKIAKRYGRDEAEATRVARRVFETWHKT
LQRTLVNLFRRLMADLVAHDPNICSFAKYIDWMCCLGIVPVIRRRRAAHSILQSKDNS
RNMLGGEFDSAACRLLTAQSIMDHGRQITDHISASMQAVTIIEYDRAIIEYNFLKRQL
RVRDAMSGERGECIVIWRPVMSDNGVVFDSPMQRICKEIVECHDLRIHAALCRLINTP
PVRVLIAKRDPEKKGVSGAQRAVEQVLGDQVENAAGSAASRLVKLIIGLKGMRHVGEI
TDTVRDYLEETNGHLLDTSSVDTSQPGFGRLNRTPNTRAADGSHSNNANIRDAFHASV
VSNINEMLEGYVTKLFSTVEGLKATNKDLVTRLSSKELELDRIRMESIQAKQARADMA
YHSQYSPNLEMLARDLKHDIIDVGANMDDDSYVANSFQSRYIPCCDTDLRRLSDLWEQ
ELMRCFKMNRVMNSQAKEIAILYSNSSITSMLAPYFFNVLNIYDIGLWVDCEDVYKSE
EDLCNSLFQKTRVHTYLQDLATFFEVDVRRAVAKLSDVFYWGRPDTLEIAPDLSYGRR
HKHRSPPTRFKLHSGRFNYRRSKRRRRSGSNYRRTDPTAYGENDDGD"

gene 12997. .13935
/gene="UL7"

CDS 12997. .13935
/gene="UL7"
/codon_start=1
/product="UL7"
/protein_id="AAG30046.1"
/db_xref="GI:11095839"
/translation="MEDDTNIVPHPLDSSCTLDALIIDALREGEGQEAIIEELIRQPM
ARMMMEIREIRGVSTRFTGISVNKLRVDQGSKRLCLTLAGADAGSEIDSHVYYQQCMT
NTAFKGFVFIVLTAAEDIVKSIGVPPSLLKFRLALFRPRDHLDFALCLLVAFLENRVG
DASDVSLYVQIQSFLKYAWTRVTSINKMRRFLCVTITWLLNTSMELGSICPFDRGYVL
PHYSIYKHLNGTWSMCANVIRAIYLQDVPSTMSQAQSALRGRDSIVLKNEGLLNGAFQ
KDWIGDVIYDWWKCDGDKLWGEEGLFGTYVGKGDNK"

gene complement(13963. .16227)
/gene="UL8"

CDS complement(13963. .16227)
/gene="UL8"
/codon_start=1
/product="UL8 DNA helicase/primase associated protein"
/protein_id="AAG30047.1"
/db_xref="GI:11095840"
/translation="MNDISAVELQGCICYATLYRLWTSSNHDEGLTALCYLLFRDVGG

NYAAHYATVNFSGKSLARAWGWNPETITDATLASMANAAATNDRLPFLNADQNTLWKA
LLAFTLSNLRHRLGYNAYYTPINMHVDKITGLVLSCTPLNSEHYKPRPGLLKTDGTIN
IDESSLLSAAMKFTNGTTLARVKLSAMSNIELADFDAKIEIQTKNDIFYREYTSRSSR
TKKTKLHVSDNAFHVAEKPLIIDGASVRTLTLVPASFDILVTTPSEFSLSALMIIYTK
WRETLFSTAGGPEKILTPIFTYIGPETNPEGEDLDYACIVGFPGWPMVKSSTASHLAI
KEVIDSYAATDGLWPMAGARTFHLLAPWRPDNHPSPTIDICHYLAESNVAQGRIATDW
GEGRLTCFFREPTLIENAAIAKLDFSAFFATLYLGLFPCHTRLRDAVLARLRRESDWL
KRPIVEFGGLLKHFREDIYKAIVGIGNLISLEIESAASTLMFAPCIYVRDGMWGAFID
KANGTIVSPEGNGNIFETLRVLCNESANRLTTSIGLRFPDGINLQMRLEGVYTHAMSW

Query Match 5.7%; Score 40.4; DB 14; Length 151102;
Best Local Similarity 61.3%; Pred. No. 1.8;
Matches 65; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

```
Qy      212 ACCATAATGCAAATACGATACGCCGAACGGCCATGTTTGCCGAGACTGACCCTATGACAT 271
            || |  | ||    ||  |||  |||||||| ||||||||||     |||       ||
Db    63566 ACAACGACGCGCGAACTTTACAGCGAACGGCAATGTTTGCCGCCGACGACGTCGCATCAC 63625

Qy      272 GGTTAAGACCAACGGTTGGCTTAAAACGTACGTTTAACCCGCGTAT 317
            || | |||||| |  |||| |||||  || || |     ||| ||||
Db    63626 GGATCAGACCATCCATTGGATTAAAGCGCACCTACCCACCGGGTAT 63671
```

Search completed: September 25, 2005, 22:51:56
Job time : 3613 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: September 25, 2005, 21:03:27 ; Search time 533 Seconds
(without alignments)
7830.052 Million cell updates/sec

Title: US-09-874-140-1
Perfect score: 705
Sequence: 1 atgacacaacccgcatcgtc.....cttcaagaagtcgtagggtg 705

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	705	100.0	705	2	AAX78759	Aax78759 VZV VP26
c	2	705	100.0	124884	5	AAH74201	Aah74201 Nucleotid
c	3	705	100.0	124884	10	ADA14878	Ada14878 Human her
c	4	705	100.0	124884	11	ADL99489	Adl99489 Varicella
c	5	703.4	99.8	125157	5	AAH74202	Aah74202 Nucleotid
c	6	38.4	5.4	2000	8	ADA71938	Ada71938 Rice gene
c	7	37.8	5.4	144411	12	ADP74214	Adp74214 Equine he
c	8	37.8	5.4	144486	12	ADP74215	Adp74215 Equine he
c	9	37.8	5.4	145444	12	ADP74213	Adp74213 Equine he
c	10	37.8	5.4	145596	12	ADP74202	Adp74202 Equine he
	11	37.6	5.3	4590	5	AAH24065	Aah24065 Yeast AOD
c	12	37	5.2	711	8	ACA43990	Aca43990 Prokaryot
	13	36.2	5.1	110000	4	AAI99682_27	Continuation (28 o
	14	36.2	5.1	110000	4	AAI99682_28	Continuation (29 o
	15	36.2	5.1	110000	4	AAI99683_27	Continuation (28 o
	16	36.2	5.1	110000	4	AAI99683_28	Continuation (29 o
c	17	35.8	5.1	1236	8	ACA38456	Aca38456 Prokaryot
c	18	35.8	5.1	1383	8	ACA40550	Aca40550 Prokaryot
c	19	35.8	5.1	110000	4	AAI99682_19	Continuation (20 o
c	20	35.8	5.1	110000	4	AAI99683_19	Continuation (20 o
	21	35.6	5.0	1222	10	ADD29744	Add29744 Mouse tum
c	22	35.6	5.0	13416	10	AAD55815	Aad55815 Micromono
	23	35.6	5.0	28874	3	AAA81505	Aaa81505 N. mening
c	24	35.6	5.0	60196	10	AAD55810	Aad55810 Micromono
	25	35.6	5.0	110000	3	AAA81490_07	Continuation (8 of
	26	35.6	5.0	110000	3	AAA81490_08	Continuation (9 of
	27	35.6	5.0	110000	4	AAI99682_22	Continuation (23 o
	28	35.6	5.0	110000	4	AAI99683_22	Continuation (23 o
	29	35.6	5.0	349980	3	AAF21608	Aaf21608 Neisseria
c	30	35.4	5.0	646	6	ABQ46950	Abq46950 Oligonucl
	31	35.4	5.0	646	6	ABQ46951	Abq46951 Oligonucl
	32	35.2	5.0	557	6	ABQ43461	Abq43461 Oligonucl
c	33	35.2	5.0	557	6	ABQ43460	Abq43460 Oligonucl
	34	34.8	4.9	1267	8	ACA23992	Aca23992 Prokaryot
c	35	34.6	4.9	11238	10	AAD55817	Aad55817 Micromono
c	36	34.6	4.9	12531	4	AAS59530	Aas59530 Propionib
c	37	34.6	4.9	12531	8	ACF64459	Acf64459 Propionib
	38	34.4	4.9	464	8	ABQ84714	Abq84714 Ostrinia
c	39	34.4	4.9	110000	4	AAI99683_39	Continuation (40 o
	40	34.2	4.9	597	4	AAF22575	Aaf22575 Human bre
	41	34.2	4.9	765	4	AAF22485	Aaf22485 Human bre
	42	34.2	4.9	54732	13	ABD33335	Abd33335 Human can
	43	34	4.8	604	6	ABQ29893	Abq29893 Oligonucl
c	44	34	4.8	604	6	ABQ29892	Abq29892 Oligonucl
	45	34	4.8	642	2	AAV23478	Aav23478 Pseudomon

ALIGNMENTS

RESULT 1
AAX78759
ID AAX78759 standard; DNA; 705 BP.
XX
AC AAX78759;
XX
DT 03-SEP-1999 (first entry)
XX
DE VZV VP26 DNA.
XX
KW VP26; immunoreactive protein; immunoassay; detection; anti-VZV; antibody;
KW ds.
XX
OS Human herpesvirus 3.
XX
PN DE19757765-A1.
XX
PD 24-JUN-1999.
XX
PF 23-DEC-1997; 97DE-01057765.
XX
PR 23-DEC-1997; 97DE-01057765.
XX
PA (DADE-) DADE BEHRING MARBURG GMBH.
XX
PI Eickmann M, Gicklhorn D, Radsak K, Hauser H, Giesendorf B;
XX
DR WPI; 1999-358931/31.
DR P-PSDB; AAY25337.
XX
PT Varicella zoster virus VP26 peptide - and corresponding nucleic acid,
PT useful for diagnosis of VZV infections.
XX
PS Claim 5; Page 5-7; 12pp; German.
XX
CC This invention describes a novel immunoreactive peptide comprising amino
CC acids 12-235 of varicella zoster virus (VZV) protein VP26, and a nucleic
CC acid encoding the peptide. Also claimed is a nucleic acid corresponding
CC to a defined DNA sequence of 705 bp given in the specification, coding
CC for amino acids 1-235 of VZV VP26. The peptide can be used in
CC immunoassays for detecting anti-VZV antibodies. The nucleic acid can be
CC used in hybridisation assays for detecting VZV
XX

```
SQ   Sequence 705 BP; 180 A; 227 C; 166 G; 132 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 705;  DB 2;  Length 705;
  Best Local Similarity   100.0%;  Pred. No. 3.5e-211;
  Matches  705;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGACACAACCCGCATCGTCTCGTGTAGTCTTTGATCCCAGCAACCCCACCACATTTTCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGACACAACCCGCATCGTCTCGTGTAGTCTTTGATCCCAGCAACCCCACCACATTTTCG 60

Qy         61 GTGGAAGCAATTGCGGCTTACACCCCCGTTGCTTTAATACGACTTTTAAACGCCAGTGGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTGGAAGCAATTGCGGCTTACACCCCCGTTGCTTTAATACGACTTTTAAACGCCAGTGGA 120

Qy        121 CCTTTGCAACCTGGTCACCGTGTGGACATCGCTGATGCCAGAAGCATTTACACCGTGGGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCTTTGCAACCTGGTCACCGTGTGGACATCGCTGATGCCAGAAGCATTTACACCGTGGGA 180

Qy        181 GCCGCGGCCAGTGCCGCGCGTGCACGCGCTAACCATAATGCAAATACGATACGCCGAACG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCCGCGGCCAGTGCCGCGCGTGCACGCGCTAACCATAATGCAAATACGATACGCCGAACG 240

Qy        241 GCCATGTTTGCCGAGACTGACCCTATGACATGGTTAAGACCAACGGTTGGCTTAAAACGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCCATGTTTGCCGAGACTGACCCTATGACATGGTTAAGACCAACGGTTGGCTTAAAACGT 300

Qy        301 ACGTTTAACCCGCGTATTATACGACCACAACCCCCAAATCCATCCATGAGTTTGGGAATC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ACGTTTAACCCGCGTATTATACGACCACAACCCCCAAATCCATCCATGAGTTTGGGAATC 360

Qy        361 TCGGGGCCTACTATATTGCCGCAAAAAACACAGAGCGCCGATCAGTCTGCTTTACAACAG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TCGGGGCCTACTATATTGCCGCAAAAAACACAGAGCGCCGATCAGTCTGCTTTACAACAG 420

Qy        421 CCCGCCGCGTTGGCGTTTTCGGGATCATCCCCGCAACACCCCCCACCTCAAACGACGTCG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CCCGCCGCGTTGGCGTTTTCGGGATCATCCCCGCAACACCCCCCACCTCAAACGACGTCG 480

Qy        481 GCATCCGTTGGACAACAGCAACACGTGGTGTCGGGGTCTTCTGGACAACAACCGCAACAG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCATCCGTTGGACAACAGCAACACGTGGTGTCGGGGTCTTCTGGACAACAACCGCAACAG 540

Qy        541 GGAGCACAGTCAAGCACTGTCCAGCCAACAACCGGATCACCGCCCGCGGCCCAAGGCGTG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GGAGCACAGTCAAGCACTGTCCAGCCAACAACCGGATCACCGCCCGCGGCCCAAGGCGTG 600

Qy        601 CCACAGTCTACCCCGCCCCCAACCCAAAATACCCCCCAGGGGGGTAAGGGACAGACCTTG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CCACAGTCTACCCCGCCCCCAACCCAAAATACCCCCCAGGGGGGTAAGGGACAGACCTTG 660

Qy        661 TCACACACGGGACAATCTGGAAACGCTTCAAGAAGTCGTAGGGTG 705
              |||||||||||||||||||||||||||||||||||||||||||||
Db        661 TCACACACGGGACAATCTGGAAACGCTTCAAGAAGTCGTAGGGTG 705


RESULT 2
AAH74201/c
ID   AAH74201 standard; DNA; 124884 BP.
XX
AC   AAH74201;
XX
DT   15-OCT-2001  (first entry)
XX
DE   Nucleotide sequence of the Varicella virus Dumas strain.
XX
KW   Dumas strain; pox vaccine; vaccine; poxvirus disease; chickenpox; ss.
XX
OS   Varicella virus.
XX
PN   WO200156600-A1.
XX
PD   09-AUG-2001.
XX
PF   31-JAN-2001; 2001WO-JP000678.
XX
PR   31-JAN-2000; 2000JP-00062734.
XX
PA   (OSAU ) UNIV OSAKA.
XX
PI   Gomi Y,  Sunamachi H,  Takahashi M,  Yamanishi K;
XX
DR   WPI; 2001-488845/53.
XX
PT   Method for checking absence of mutation at specific positions of
PT   varicella virus genome for quality control of attenuated live varicella
PT   vaccine.
XX
PS   Claim 1; Page 61-105; 158pp; Japanese.
XX
CC   The present sequence represents the nucleotide sequence of Varicella
CC   virus, Dumas strain. The specification describes a method for the quality
CC   control of an attenuated pox vaccine, characterised in that the DNA
CC   sequence of specific parts of the viral genome in a sample is determined
CC   and proved to be conserved rather than mutated in comparison with the
CC   same parts of a reference viral genome. The method is useful for quality
CC   control of vaccines for use in the prevention of poxvirus diseases such
CC   as chickenpox
XX
SQ   Sequence 124884 BP; 33789 A; 29295 C; 28177 G; 33623 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 705;  DB 5;  Length 124884;
  Best Local Similarity   100.0%;  Pred. No. 3.3e-210;
  Matches  705;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGACACAACCCGCATCGTCTCGTGTAGTCTTTGATCCCAGCAACCCCACCACATTTTCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      43138 ATGACACAACCCGCATCGTCTCGTGTAGTCTTTGATCCCAGCAACCCCACCACATTTTCG 43079

Qy         61 GTGGAAGCAATTGCGGCTTACACCCCCGTTGCTTTAATACGACTTTTAAACGCCAGTGGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      43078 GTGGAAGCAATTGCGGCTTACACCCCCGTTGCTTTAATACGACTTTTAAACGCCAGTGGA 43019

Qy        121 CCTTTGCAACCTGGTCACCGTGTGGACATCGCTGATGCCAGAAGCATTTACACCGTGGGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      43018 CCTTTGCAACCTGGTCACCGTGTGGACATCGCTGATGCCAGAAGCATTTACACCGTGGGA 42959

Qy        181 GCCGCGGCCAGTGCCGCGCGTGCACGCGCTAACCATAATGCAAATACGATACGCCGAACG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42958 GCCGCGGCCAGTGCCGCGCGTGCACGCGCTAACCATAATGCAAATACGATACGCCGAACG 42899

Qy        241 GCCATGTTTGCCGAGACTGACCCTATGACATGGTTAAGACCAACGGTTGGCTTAAAACGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42898 GCCATGTTTGCCGAGACTGACCCTATGACATGGTTAAGACCAACGGTTGGCTTAAAACGT 42839

Qy        301 ACGTTTAACCCGCGTATTATACGACCACAACCCCCAAATCCATCCATGAGTTTGGGAATC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42838 ACGTTTAACCCGCGTATTATACGACCACAACCCCCAAATCCATCCATGAGTTTGGGAATC 42779

Qy        361 TCGGGGCCTACTATATTGCCGCAAAAAACACAGAGCGCCGATCAGTCTGCTTTACAACAG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42778 TCGGGGCCTACTATATTGCCGCAAAAAACACAGAGCGCCGATCAGTCTGCTTTACAACAG 42719

Qy        421 CCCGCCGCGTTGGCGTTTTCGGGATCATCCCCGCAACACCCCCCACCTCAAACGACGTCG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42718 CCCGCCGCGTTGGCGTTTTCGGGATCATCCCCGCAACACCCCCCACCTCAAACGACGTCG 42659

Qy        481 GCATCCGTTGGACAACAGCAACACGTGGTGTCGGGGTCTTCTGGACAACAACCGCAACAG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42658 GCATCCGTTGGACAACAGCAACACGTGGTGTCGGGGTCTTCTGGACAACAACCGCAACAG 42599

Qy        541 GGAGCACAGTCAAGCACTGTCCAGCCAACAACCGGATCACCGCCCGCGGCCCAAGGCGTG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42598 GGAGCACAGTCAAGCACTGTCCAGCCAACAACCGGATCACCGCCCGCGGCCCAAGGCGTG 42539

Qy        601 CCACAGTCTACCCCGCCCCCAACCCAAAATACCCCCCAGGGGGGTAAGGGACAGACCTTG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db      42538 CCACAGTCTACCCCGCCCCCAACCCAAAATACCCCCCAGGGGGGTAAGGGACAGACCTTG 42479

Qy        661 TCACACACGGGACAATCTGGAAACGCTTCAAGAAGTCGTAGGGTG 705
              |||||||||||||||||||||||||||||||||||||||||||||
Db      42478 TCACACACGGGACAATCTGGAAACGCTTCAAGAAGTCGTAGGGTG 42434


RESULT 3
ADA14878/c
ID   ADA14878 standard; DNA; 124884 BP.
XX
AC   ADA14878;
XX
DT   06-NOV-2003  (first entry)
XX
DE   Human herpesvirus 3 DNA.
XX
KW   ds; varicella zoster virus; VZV; human herpesvirus 3; vaccine;
KW   chicken pox; shingles.
XX
OS   Human herpesvirus 3.
XX
PN   US6528066-B1.
XX
PD   04-MAR-2003.
XX
PF   14-SEP-2000; 2000US-00661596.
XX
PR   14-SEP-1999;   99US-0153779P.
XX
PA   (IOWA ) UNIV IOWA RES FOUND.
XX
PI   Grose CF,  Santos R;
XX
DR   WPI; 2003-361635/34.
XX
PT   New isolated polynucleotide isolated from varicella zoster virus, useful
PT   for diagnosing a disease, for instance chicken pox and shingles, caused
PT   by the virus.
XX
PS   Disclosure; Col 61-168; 98pp; English.
XX
CC   The invention relates to an isolated polynucleotide isolated from human
CC   herpesvirus 3 (varicella zoster virus (VZV)). The polynucleotide is
CC   useful for creating kits for detecting antibodies that specifically bind
CC   to VZV polypeptide, for detecting the presence of VZV in an animal, for a
CC   vaccine composition that includes a modified attenuated VZV, for
CC   producing a modified attenuated VZV, for detecting VZV, such as VZV-MSP,
CC   having a single nucleotide polymorphism in open reading frame (ORF) 68,
CC   or for diagnosing a disease, for instance chicken pox and shingles,
CC   caused by VZV. The present sequence represents human herpesvirus 3 DNA.
XX
SQ   Sequence 124884 BP; 33789 A; 29295 C; 28177 G; 33623 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 705;  DB 10;  Length 124884;
  Best Local Similarity   100.0%;  Pred. No. 3.3e-210;
  Matches  705;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGACACAACCCGCATCGTCTCGTGTAGTCTTTGATCCCAGCAACCCCACCACATTTTCG 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      43138 ATGACACAACCCGCATCGTCTCGTGTAGTCTTTGATCCCAGCAACCCCACCACATTTTCG 43079

Qy         61 GTGGAAGCAATTGCGGCTTACACCCCCGTTGCTTTAATACGACTTTTAAACGCCAGTGGA 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      43078 GTGGAAGCAATTGCGGCTTACACCCCCGTTGCTTTAATACGACTTTTAAACGCCAGTGGA 43019

Qy        121 CCTTTGCAACCTGGTCACCGTGTGGACATCGCTGATGCCAGAAGCATTTACACCGTGGGA 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      43018 CCTTTGCAACCTGGTCACCGTGTGGACATCGCTGATGCCAGAAGCATTTACACCGTGGGA 42959

Qy        181 GCCGCGGCCAGTGCCGCGCGTGCACGCGCTAACCATAATGCAAATACGATACGCCGAACG 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42958 GCCGCGGCCAGTGCCGCGCGTGCACGCGCTAACCATAATGCAAATACGATACGCCGAACG 42899

Qy        241 GCCATGTTTGCCGAGACTGACCCTATGACATGGTTAAGACCAACGGTTGGCTTAAAACGT 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42898 GCCATGTTTGCCGAGACTGACCCTATGACATGGTTAAGACCAACGGTTGGCTTAAAACGT 42839

Qy        301 ACGTTTAACCCGCGTATTATACGACCACAACCCCCAAATCCATCCATGAGTTTGGGAATC 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42838 ACGTTTAACCCGCGTATTATACGACCACAACCCCCAAATCCATCCATGAGTTTGGGAATC 42779

Qy        361 TCGGGGCCTACTATATTGCCGCAAAAAACACAGAGCGCCGATCAGTCTGCTTTACAACAG 420
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42778 TCGGGGCCTACTATATTGCCGCAAAAAACACAGAGCGCCGATCAGTCTGCTTTACAACAG 42719

Qy        421 CCCGCCGCGTTGGCGTTTTCGGGATCATCCCCGCAACACCCCCCACCTCAAACGACGTCG 480
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42718 CCCGCCGCGTTGGCGTTTTCGGGATCATCCCCGCAACACCCCCCACCTCAAACGACGTCG 42659

Qy        481 GCATCCGTTGGACAACAGCAACACGTGGTGTCGGGGTCTTCTGGACAACAACCGCAACAG 540
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42658 GCATCCGTTGGACAACAGCAACACGTGGTGTCGGGGTCTTCTGGACAACAACCGCAACAG 42599

Qy        541 GGAGCACAGTCAAGCACTGTCCAGCCAACAACCGGATCACCGCCCGCGGCCCAAGGCGTG 600
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42598 GGAGCACAGTCAAGCACTGTCCAGCCAACAACCGGATCACCGCCCGCGGCCCAAGGCGTG 42539

Qy        601 CCACAGTCTACCCCGCCCCCAACCCAAAATACCCCCCAGGGGGGTAAGGGACAGACCTTG 660
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42538 CCACAGTCTACCCCGCCCCCAACCCAAAATACCCCCCAGGGGGGTAAGGGACAGACCTTG 42479

Qy        661 TCACACACGGGACAATCTGGAAACGCTTCAAGAAGTCGTAGGGTG 705
              |||||||||||||||||||||||||||||||||||||||||||||
Db      42478 TCACACACGGGACAATCTGGAAACGCTTCAAGAAGTCGTAGGGTG 42434


RESULT 4
ADL99489/c
ID   ADL99489 standard; DNA; 124884 BP.
XX
AC   ADL99489;
XX
DT   20-MAY-2004  (first entry)
XX
DE   Varicella-Zoster virus complete genome.
XX
KW   Varicella-Zoster virus; ds; ORF68; ORF37; vaccine; VZV infection; SNP;
KW   single nucleotide polymorphism.
XX
OS   Human herpesvirus 3.
XX
FH   Key             Location/Qualifiers
FT   variation       replace(66879,T)
FT                   /*tag=  a
FT                   /standard_name= "Single nucleotide polymorphism"
FT                   /note= "Claimed in claim 60"
FT   variation       replace(102203,B)
FT                   /*tag=  b
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(102575,B)
FT                   /*tag=  c
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(102617,D)
FT                   /*tag=  d
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(102969,B)
FT                   /*tag=  e
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(104898,B)
FT                   /*tag=  f
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(116255,A)
FT                   /*tag=  g
FT                   /standard_name= "Single nucleotide polymorphism"
FT                   /note= "Claimed in claim 62"
```

XX
PN US2003166168-A1.
XX
PD 04-SEP-2003.
XX
PF 06-NOV-2002; 2002US-00288823.
XX
PR 14-SEP-1999; 99US-0153779P.
PR 14-SEP-2000; 2000US-00661596.
XX
PA (IOWA) UNIV IOWA RES FOUND.
XX
PI Grose CF, Santos R;
XX
DR WPI; 2003-898077/82.
DR GENBANK; X04370.
XX
PT Detecting antibodies that bind to a varicella zoster polypeptide with
PT polymorphisms in ORF37 or ORF68 comprises contacting sample with viral
PT preparation having polymorphism and detecting presence of complex.
XX
PS Claim 17; Page; 33pp; English.
XX
CC The invention relates to detecting antibodies that specifically bind to a
CC varicella zoster polypeptide, comprising contacting a sample having an
CC antibody with a preparation comprising a varicella zoster polypeptide
CC having a polymorphism to form a mixture and detecting the presence or
CC absence of the complex. The polypeptide is encoded by a polymorphism of
CC open reading frame 37 (ORF37) and/or ORF68. Also included are a kit for
CC detecting antibodies that specifically bind to a varicella zoster
CC polypeptide, a method for detecting the presence of a varicella zoster
CC virus in an animal, a method for diagnosing a disease caused by varicella
CC zoster virus, a method for detecting a varicella zoster virus having a
CC single nucleotide polymorphism (SNP) in ORF68, a vaccine composition
CC comprising a modified attenuated varicella zoster virus having the ATCC
CC designation VR-795 where the nucleotide sequence of the virus has been
CC modified to contain an SNP, a method for producing a modified attenuated
CC varicella zoster virus having the ATCC designation VR-795, an isolated
CC polynucleotide having the nucleotide sequence of nucleotides 66074 to
CC 68599 of GenBank X04370 (with the proviso that nucleotide 66879 is a
CC thymine) or nucleotides 115808 to 117679 of GenBank Accession X04370
CC (with the proviso that nucleotide 116255 is an adenine), an isolated
CC polypeptide encoded by the polynucleotide and a virus having the
CC designation VZV-MSP, VZV-VSD, VZV-VIA and/or VZV-Iceland. The methods and
CC compositions of the present invention are useful for diagnosing,
CC detecting and/or treating varicella zoster virus (VZV) infection. The
CC present sequence is the VZV complete genome.
XX
SQ Sequence 124884 BP; 33789 A; 29295 C; 28177 G; 33623 T; 0 U; 0 Other;

Query Match 100.0%; Score 705; DB 11; Length 124884;
Best Local Similarity 100.0%; Pred. No. 3.3e-210;
Matches 705; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ATGACACAACCCGCATCGTCTCGTGTAGTCTTTGATCCCAGCAACCCCACCACATTTTCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      43138 ATGACACAACCCGCATCGTCTCGTGTAGTCTTTGATCCCAGCAACCCCACCACATTTTCG 43079

Qy         61 GTGGAAGCAATTGCGGCTTACACCCCCGTTGCTTTAATACGACTTTTAAACGCCAGTGGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      43078 GTGGAAGCAATTGCGGCTTACACCCCCGTTGCTTTAATACGACTTTTAAACGCCAGTGGA 43019

Qy        121 CCTTTGCAACCTGGTCACCGTGTGGACATCGCTGATGCCAGAAGCATTTACACCGTGGGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      43018 CCTTTGCAACCTGGTCACCGTGTGGACATCGCTGATGCCAGAAGCATTTACACCGTGGGA 42959

Qy        181 GCCGCGGCCAGTGCCGCGCGTGCACGCGCTAACCATAATGCAAATACGATACGCCGAACG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42958 GCCGCGGCCAGTGCCGCGCGTGCACGCGCTAACCATAATGCAAATACGATACGCCGAACG 42899

Qy        241 GCCATGTTTGCCGAGACTGACCCTATGACATGGTTAAGACCAACGGTTGGCTTAAAACGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42898 GCCATGTTTGCCGAGACTGACCCTATGACATGGTTAAGACCAACGGTTGGCTTAAAACGT 42839

Qy        301 ACGTTTAACCCGCGTATTATACGACCACAACCCCCAAATCCATCCATGAGTTTGGGAATC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42838 ACGTTTAACCCGCGTATTATACGACCACAACCCCCAAATCCATCCATGAGTTTGGGAATC 42779

Qy        361 TCGGGGCCTACTATATTGCCGCAAAAAACACAGAGCGCCGATCAGTCTGCTTTACAACAG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42778 TCGGGGCCTACTATATTGCCGCAAAAAACACAGAGCGCCGATCAGTCTGCTTTACAACAG 42719

Qy        421 CCCGCCGCGTTGGCGTTTTCGGGATCATCCCCGCAACACCCCCCACCTCAAACGACGTCG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42718 CCCGCCGCGTTGGCGTTTTCGGGATCATCCCCGCAACACCCCCCACCTCAAACGACGTCG 42659

Qy        481 GCATCCGTTGGACAACAGCAACACGTGGTGTCGGGGTCTTCTGGACAACAACCGCAACAG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42658 GCATCCGTTGGACAACAGCAACACGTGGTGTCGGGGTCTTCTGGACAACAACCGCAACAG 42599

Qy        541 GGAGCACAGTCAAGCACTGTCCAGCCAACAACCGGATCACCGCCCGCGGCCCAAGGCGTG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42598 GGAGCACAGTCAAGCACTGTCCAGCCAACAACCGGATCACCGCCCGCGGCCCAAGGCGTG 42539

Qy        601 CCACAGTCTACCCCGCCCCCAACCCAAAATACCCCCCAGGGGGGTAAGGGACAGACCTTG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42538 CCACAGTCTACCCCGCCCCCAACCCAAAATACCCCCCAGGGGGGTAAGGGACAGACCTTG 42479

Qy        661 TCACACACGGGACAATCTGGAAACGCTTCAAGAAGTCGTAGGGTG 705
              |||||||||||||||||||||||||||||||||||||||||||||
Db      42478 TCACACACGGGACAATCTGGAAACGCTTCAAGAAGTCGTAGGGTG 42434
```

RESULT 5
AAH74202/c
ID AAH74202 standard; DNA; 125157 BP.
XX
AC AAH74202;
XX
DT 15-OCT-2001 (first entry)
XX
DE Nucleotide sequence of the Varicella virus Oka strain.
XX
KW Oka strain; pox vaccine; vaccine; poxvirus disease; chickenpox; ss.
XX
OS Varicella virus.
XX
PN WO200156600-A1.
XX
PD 09-AUG-2001.
XX
PF 31-JAN-2001; 2001WO-JP000678.
XX
PR 31-JAN-2000; 2000JP-00062734.
XX
PA (OSAU) UNIV OSAKA.
XX
PI Gomi Y, Sunamachi H, Takahashi M, Yamanishi K;
XX
DR WPI; 2001-488845/53.
XX
PT Method for checking absence of mutation at specific positions of
PT varicella virus genome for quality control of attenuated live varicella
PT vaccine.
XX
PS Disclosure; Page 106-150; 158pp; Japanese.
XX
CC The present sequence represents the nucleotide sequence of Varicella
CC virus, Oka strain. The specification describes a method for the quality
CC control of an attenuated pox vaccine, characterised in that the DNA
CC sequence of specific parts of the viral genome in a sample is determined
CC and proved to be conserved rather than mutated in comparison with the
CC same parts of a reference viral genome. The method is useful for quality
CC control of vaccines for use in the prevention of poxvirus diseases such
CC as chickenpox
XX

```
SQ   Sequence 125157 BP; 33776 A; 29439 C; 28285 G; 33624 T; 0 U; 33 Other;

  Query Match            99.8%;  Score 703.4;  DB 5;  Length 125157;
  Best Local Similarity  99.9%;  Pred. No. 1e-209;
  Matches  704;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 ATGACACAACCCGCATCGTCTCGTGTAGTCTTTGATCCCAGCAACCCCACCACATTTTCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      43113 ATGACACAACCCGCATCGTCTCGTGTAGTCTTTGATCCCAGCAACCCCACCACATTTTCG 43054

Qy         61 GTGGAAGCAATTGCGGCTTACACCCCCGTTGCTTTAATACGACTTTTAAACGCCAGTGGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      43053 GTGGAAGCAATTGCGGCTTACACCCCCGTTGCTTTAATACGACTTTTAAACGCCAGTGGA 42994

Qy        121 CCTTTGCAACCTGGTCACCGTGTGGACATCGCTGATGCCAGAAGCATTTACACCGTGGGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42993 CCTTTGCAACCTGGTCACCGTGTGGACATCGCTGATGCCAGAAGCATTTACACCGTGGGA 42934

Qy        181 GCCGCGGCCAGTGCCGCGCGTGCACGCGCTAACCATAATGCAAATACGATACGCCGAACG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42933 GCCGCGGCCAGTGCCGCGCGTGCACGCGCTAACCATAATGCAAATACGATACGCCGAACG 42874

Qy        241 GCCATGTTTGCCGAGACTGACCCTATGACATGGTTAAGACCAACGGTTGGCTTAAAACGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42873 GCCATGTTTGCCGAGACTGACCCTATGACATGGTTAAGACCAACGGTTGGCTTAAAACGT 42814

Qy        301 ACGTTTAACCCGCGTATTATACGACCACAACCCCCAAATCCATCCATGAGTTTGGGAATC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42813 ACGTTTAACCCGCGTATTATACGACCACAACCCCCAAATCCATCCATGAGTTTGGGAATC 42754

Qy        361 TCGGGGCCTACTATATTGCCGCAAAAAACACAGAGCGCCGATCAGTCTGCTTTACAACAG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42753 TCGGGGCCTACTATATTGCCGCAAAAAACACAGAGCGCCGATCAGTCTGCTTTACAACAG 42694

Qy        421 CCCGCCGCGTTGGCGTTTTCGGGATCATCCCCGCAACACCCCCCACCTCAAACGACGTCG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42693 CCCGCCGCGTTGGCGTTTTCGGGATCATCCCCGCAACACCCCCCACCTCAAACGACGTCG 42634

Qy        481 GCATCCGTTGGACAACAGCAACACGTGGTGTCGGGGTCTTCTGGACAACAACCGCAACAG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42633 GCATCCGTTGGACAACAGCAACACGTGGTGTCGGGGTCTTCTGGACAACAACCGCAACAG 42574

Qy        541 GGAGCACAGTCAAGCACTGTCCAGCCAACAACCGGATCACCGCCCGCGGCCCAAGGCGTG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42573 GGAGCACAGTCAAGCACTGTCCAGCCAACAACCGGATCACCGCCCGCGGCCCAAGGCGTG 42514

Qy        601 CCACAGTCTACCCCGCCCCCAACCCAAAATACCCCCCAGGGGGGTAAGGGACAGACCTTG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42513 CCACAGTCTACCCCGCCCCCAACCCAAAATACCCCCCAGGGGGGTAAGGGACAGACCTTG 42454

Qy        661 TCACACACGGGACAATCTGGAAACGCTTCAAGAAGTCGTAGGGTG 705
              || |||||||||||||||||||||||||||||||||||||||||||
Db      42453 TCCCACACGGGACAATCTGGAAACGCTTCAAGAAGTCGTAGGGTG 42409


RESULT 6
ADA71938/c
ID   ADA71938 standard; DNA; 2000 BP.
XX
AC   ADA71938;
XX
DT   20-NOV-2003  (first entry)
XX
DE   Rice gene, SEQ ID 5263.
XX
KW   Plant; bacterial infection; fungal infection; viral infection; rice;
KW   gene; ds.
XX
OS   Oryza sativa.
XX
PN   WO2003000898-A1.
XX
PD   03-JAN-2003.
XX
PF   22-JUN-2001; 2001WO-IB001105.
XX
PR   22-JUN-2001; 2001WO-IB001105.
XX
PA   (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX
PI   Chang H,  Chen W,  Cooper B,  Glazebrook J,  Goff SA,  Hou Y;
PI   Katagiri F,  Quan S,  Tao Y,  Whitham S,  Xie Z,  Zhu T,  Zou G;
XX
DR   WPI; 2003-175290/17.
XX
PT   Identifying at least one gene involved in plant resistance or response to
PT   pathogenic infection for conferring resistance or tolerance to a plant to
PT   bacterial, fungal or viral infection by determining or detecting plant
PT   gene expression.
XX
PS   Claim 27; SEQ ID NO 5263; 899pp; English.
XX
CC   The present invention relates to a method (M1) for identifying genes
CC   involved in plant resistance or response to pathogenic infection. M1
CC   comprises identifying a gene whose expression is significantly altered in
CC   the incompatible interaction of plant gene expression relative to
CC   expression of the gene in an uninfected plant, in a mutant plant that
CC   does not express a gene associated with response to pathogenic infection,
CC   or in a corresponding incompatible or compatible interaction. (M1) is
CC   useful for conferring resistance to resistance or tolerance to a plant to
CC   bacterial, fungal or viral infection. The present sequence was used to
CC   illustrate the invention.
XX
SQ   Sequence 2000 BP; 336 A; 265 C; 284 G; 363 T; 0 U; 752 Other;

  Query Match             5.4%;  Score 38.4;  DB 8;  Length 2000;
  Best Local Similarity  10.8%;  Pred. No. 0.41;
  Matches   63;  Conservative  268;  Mismatches  249;  Indels    6;  Gaps    3;

Qy        118 GGACCTTTGCAACCTGGTCACCGTGTGGACATCGCTGATGCCAGAAGCATTTACACCGTG 177
              |:: || : ::  :|::  :    | :::::    : ::: ::: ::   |::  :|
Db        602 GKMTCTMYTSMKGSTRRSKMGRWSGMSRMYMRWWKKMRKRKYMRYMKWKCTWRRCMCYRW 543

Qy        178 GGAGCCGCGGCCAGTGCCGCGCGTGCACGCGCTAACCATAATGCAAATACGATACGCCGA 237
              |   :  ::: :: || :: ::  : ::    :: :|::  :|:  :: : : ::|:|:|
Db        542 GYTMYTTSRSRMMYTGRYKARYTSKRRYMWYKYRKYCWYYYYGM-YMKCSYMMRYGYCKA 484

Qy        238 ACGGCCATGTTTGCCGAGACTGACCCTATGACATGGTTAAGACCAACGGTTGGCTTAAAA 297
              : |:|  ::  ::|: ::: : :::: ::: :  : ::   |: ::   :||:   ::
Db        483 CKKCCYAMCWKAAYSGMMMYWYRKYSKWMRMSTKYMWSMWYKKCRSMKYGAKGCYGCKMW 424

Qy        298 CGTACGTTTAACCCGCGTATTATACGACCAC--AACCCCCAAATCCATCCATGAGTTTGG 355
              :|: ::  ::|: :    : : ::::: ::::::: |  :::| : : |::::
Db        423 TYCSYGYMKWYTYMGSYKYSRCYKYMRMYMYKGWMYMMYYSAYSSMMTWYYYYAKYWKYW 364

Qy        356 GAATCTCGGGGCCTACTATATTGCCGCAAAAAACACAGAGCGCCGATCAGTCTGCTTTAC 415
              :  |::  | ::  :   :: :::|::  ::  ::: : ::::: :   :  |::  |
Db        363 YKRRGTMSWYGKSYKKKYCTWWCYMKCMRCYRWRKMMRKKTKYSKRCYCWRYATCYWCCC 304

Qy        416 AACAGCCCGCCGCGTTGGCGTTTTCGGGATCATCCCCGCAACACCCCCCACCTCAAACGA 475
              : :| ::: : ::|    |:::  : :::|: :: :::   :   ::::::  |: :|
Db        303 YRKRGWYSRRSMMRT---AGKWKMRSWSRWCRSYSWYKMYKKMWKKSYYMSYGWARSSGT 247

Qy        476 CGTCGGCATCCGTTGGACAACAGCAACACGTGGTGTCGGGGTCTTCTGGACAACAACCGC 535
              : :   :|: |:  :: | ::: :: :  :::: :     | :| ::    :  :|::
Db        246 WSRSAAKRTYKGYSTSRRAKMMRACRMYSACRRYSRTSYYCGCSYCGSSKWKYMSKSCSM 187

Qy        536 AACAGGGAGCACAGTCAAGCACTGTCCAGCCAACAACCGGATCACCGCCCGCGGCCCAAG 595
              ::  ::  :| :  :: | |::   ::: :         |  :   :::||:  :||
Db        186 RMTCSSWCSCCYTCYYGAMCWSCCMSMMYMGSCGCYTRGWKWRSKYSMCCKKYCSCCTKY 127

Qy        596 GCGTGCCACAGTCTACCCCGCCCCCAACCCAAAATACCCCCCAGGGGGGTAAGGGACAGA 655
              : ||:::: :::: :::  ::|:| :::: ::  : |:|:   ::  : |  :  : ::
Db        126 CSYTGYYRYCKWYKYSYYKCYYCYCYWYMSYMRYMMKCMCSRSCSSWMSCAYCSTSSTSR 67

Qy        656 CCTTGTCACACACGGGACAATCTGGAAACGCTTCAAGAAGTCGTAG 701
```

```
              :   : | : :||:: : : ::: : ::: |     ::| :::
Db         66 WMSMYYAAKMGMCGSSGMYRMSKSCKMYSKYSSCKYTGSKKCTKRK 21
```

RESULT 7
ADP74214/c
ID ADP74214 standard; DNA; 144411 BP.
XX
AC ADP74214;
XX
DT 26-AUG-2004 (first entry)
XX
DE Equine herpesvirus 4 genome gM deletion mutant #2.
XX
KW virucide; vaccine; Equine Herpes Virus; EHV; gM; vaccine;
KW EHV-associated condition; Equine herpesvirus 4; EHV4; ds; mutant; mutein.
XX
OS Equine herpesvirus 4.
OS Synthetic.
XX
PN US2004109873-A1.
XX
PD 10-JUN-2004.
XX
PF 21-JUL-2003; 2003US-00624149.
XX
PR 19-JUL-2002; 2002DE-01033064.
PR 14-AUG-2002; 2002US-0403282P.
PR 11-APR-2003; 2003DE-01017008.
XX
PA (BOEH) BOEHRINGER INGELHEIM VETMEDICA GMBH.
XX
PI Neubauer A, Ziegler C;
XX
DR WPI; 2004-440311/41.
XX
PT New recombinant equine herpes (EHV) virus free of heterologous elements,
PT and where protein gM has been deleted, useful as a vaccine for treating
PT or preventing EHV infections.
XX
PS Claim 18; Page; 156pp; English.
XX
CC The invention describes a new recombinant Equine Herpes Virus (EHV) where
CC the protein gM is absent, and the EHV is free of heterologous elements.
CC Also described are: a nucleic acid coding for an EHV defined above; a
CC vaccine preparation comprising the EHV or nucleic acid; obtaining a
CC recombinant EHV; and a cell line for use in the method, where the gene
CC encoding the protein gM is transfected into the cell line, and the cell
CC line expresses gM. The vaccine comprising the EHV or nucleic acid
CC encoding EHV is useful for treating and/or preventing EHV-associated
CC condition, and for monitoring the therapeutic success. The recombinant
CC EHV is useful as a vaccine against EHV infections. This sequence
CC represents an Equine herpesvirus 4 genome in which nucleotides 92681-
CC 93865, comprising regions of the gM gene, have been deleted.
XX
SQ Sequence 144411 BP; 35893 A; 36645 C; 36246 G; 35627 T; 0 U; 0 Other;

  Query Match   5.4%; Score 37.8; DB 12; Length 144411;
  Best Local Similarity   61.9%; Pred. No. 4;
  Matches   60; Conservative   0; Mismatches   37; Indels   0; Gaps   0;

```
Qy        230 TACGCCGAACGGCCATGTTTGCCGAGACTGACCCTATGACATGGTTAAGACCAACGGTTG 289
              | || ||| | ||||||||||  | |  ||| |    || |||    | ||||   | |
Db      46820 TTCGACGATCTGCCATGTTTGCGCACAACGACGCATCTACCTGGGCCCGGCCAAATATAG 46761

Qy        290 GCTTAAAACGTACGTTTAACCCGCGTATTATACGACC 326
              |  ||||||| || ||    |||||  |||| |  ||
Db      46760 GACTAAAACGCACCTTCCCGCCGCGGTTTATGCAGCC 46724
```

RESULT 8
ADP74215/c
ID ADP74215 standard; DNA; 144486 BP.
XX
AC ADP74215;
XX
DT 26-AUG-2004 (first entry)
XX
DE Equine herpesvirus 4 genome gM deletion mutant #3.
XX
KW virucide; vaccine; Equine Herpes Virus; EHV; gM; vaccine;
KW EHV-associated condition; Equine herpesvirus 4; EHV4; ds; mutant; mutein.
XX
OS Equine herpesvirus 4.
OS Synthetic.
XX
PN US2004109873-A1.
XX
PD 10-JUN-2004.
XX
PF 21-JUL-2003; 2003US-00624149.
XX
PR 19-JUL-2002; 2002DE-01033064.
PR 14-AUG-2002; 2002US-0403282P.
PR 11-APR-2003; 2003DE-01017008.
XX
PA (BOEH) BOEHRINGER INGELHEIM VETMEDICA GMBH.
XX
PI Neubauer A, Ziegler C;
XX
DR WPI; 2004-440311/41.
XX
PT New recombinant equine herpes (EHV) virus free of heterologous elements,
PT and where protein gM has been deleted, useful as a vaccine for treating
PT or preventing EHV infections.
XX
PS Claim 19; Page; 156pp; English.
XX
CC The invention describes a new recombinant Equine Herpes Virus (EHV) where
CC the protein gM is absent, and the EHV is free of heterologous elements.
CC Also described are: a nucleic acid coding for an EHV defined above; a
CC vaccine preparation comprising the EHV or nucleic acid; obtaining a
CC recombinant EHV; and a cell line for use in the method, where the gene
CC encoding the protein gM is transfected into the cell line, and the cell
CC line expresses gM. The vaccine comprising the EHV or nucleic acid
CC encoding EHV is useful for treating and/or preventing EHV-associated
CC condition, and for monitoring the therapeutic success. The recombinant
CC EHV is useful as a vaccine against EHV infections. This sequence
CC represents an Equine herpesvirus 4 genome in which nucleotides 92715-
CC 93824, comprising regions of the gM gene, have been deleted.
XX
SQ Sequence 144486 BP; 35916 A; 36662 C; 36259 G; 35649 T; 0 U; 0 Other;

  Query Match   5.4%; Score 37.8; DB 12; Length 144486;
  Best Local Similarity   61.9%; Pred. No. 4;
  Matches   60; Conservative   0; Mismatches   37; Indels   0; Gaps   0;

```
Qy        230 TACGCCGAACGGCCATGTTTGCCGAGACTGACCCTATGACATGGTTAAGACCAACGGTTG 289
              | || ||| | ||||||||||  | |  ||| |    || |||    | ||||   | |
Db      46820 TTCGACGATCTGCCATGTTTGCGCACAACGACGCATCTACCTGGGCCCGGCCAAATATAG 46761

Qy        290 GCTTAAAACGTACGTTTAACCCGCGTATTATACGACC 326
              |  ||||||| || ||    |||||  |||| |  ||
Db      46760 GACTAAAACGCACCTTCCCGCCGCGGTTTATGCAGCC 46724
```

RESULT 9
ADP74213/c
ID ADP74213 standard; DNA; 145444 BP.
XX
AC ADP74213;
XX
DT 26-AUG-2004 (first entry)
XX
DE Equine herpesvirus 4 genome gM deletion mutant #1.
XX

KW virucide; vaccine; Equine Herpes Virus; EHV; gM; vaccine;
KW EHV-associated condition; Equine herpesvirus 4; EHV4; ds; mutant; mutein.
XX
OS Equine herpesvirus 4.
OS Synthetic.
XX
PN US2004109873-A1.
XX
PD 10-JUN-2004.
XX
PF 21-JUL-2003; 2003US-00624149.
XX
PR 19-JUL-2002; 2002DE-01033064.
PR 14-AUG-2002; 2002US-0403282P.
PR 11-APR-2003; 2003DE-01017008.
XX
PA (BOEH) BOEHRINGER INGELHEIM VETMEDICA GMBH.
XX
PI Neubauer A, Ziegler C;
XX
DR WPI; 2004-440311/41.
XX
PT New recombinant equine herpes (EHV) virus free of heterologous elements,
PT and where protein gM has been deleted, useful as a vaccine for treating
PT or preventing EHV infections.
XX
PS Claim 17; Page; 156pp; English.
XX
CC The invention describes a new recombinant Equine Herpes Virus (EHV) where
CC the protein gM is absent, and the EHV is free of heterologous elements.
CC Also described are: a nucleic acid coding for an EHV defined above; a
CC vaccine preparation comprising the EHV or nucleic acid; obtaining a
CC recombinant EHV; and a cell line for use in the method, where the gene
CC encoding the protein gM is transfected into the cell line, and the cell
CC line expresses gM. The vaccine comprising the EHV or nucleic acid
CC encoding EHV is useful for treating and/or preventing EHV-associated
CC condition, and for monitoring the therapeutic success. The recombinant
CC EHV is useful as a vaccine against EHV infections. This sequence
CC represents an Equine herpesvirus 4 genome in which nucleotides 92681-
CC 92731 and 93765-93865, comprising regions of the gM gene, have been
CC deleted.
XX
SQ Sequence 145444 BP; 36178 A; 36893 C; 36497 G; 35876 T; 0 U; 0 Other;

Query Match 5.4%; Score 37.8; DB 12; Length 145444;
Best Local Similarity 61.9%; Pred. No. 4;
Matches 60; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

Qy 230 TACGCCGAACGGCCATGTTTGCCGAGACTGACCCTATGACATGGTTAAGACCAACGGTTG 289
| || ||| | |||||||||| | | ||| | || ||| | |||| | |
Db 46820 TTCGACGATCTGCCATGTTTGCGCACAACGACGCATCTACCTGGGCCCGGCCAAATATAG 46761

Qy 290 GCTTAAAACGTACGTTTAACCCGCGTATTATACGACC 326
| ||||||| || || ||||| |||| | ||
Db 46760 GACTAAAACGCACCTTCCCGCCGCGGTTTATGCAGCC 46724

RESULT 10
ADP74202/c
ID ADP74202 standard; DNA; 145596 BP.
XX
AC ADP74202;
XX
DT 26-AUG-2004 (first entry)
XX
DE Equine herpesvirus 4 genome seqid 2.
XX
KW virucide; vaccine; Equine Herpes Virus; EHV; gM; vaccine;
KW EHV-associated condition; Equine herpesvirus 4; EHV4; ds.
XX
OS Equine herpesvirus 4.
XX
PN US2004109873-A1.
XX
PD 10-JUN-2004.
XX
PF 21-JUL-2003; 2003US-00624149.
XX
PR 19-JUL-2002; 2002DE-01033064.
PR 14-AUG-2002; 2002US-0403282P.
PR 11-APR-2003; 2003DE-01017008.
XX
PA (BOEH) BOEHRINGER INGELHEIM VETMEDICA GMBH.
XX
PI Neubauer A, Ziegler C;
XX
DR WPI; 2004-440311/41.
XX
PT New recombinant equine herpes (EHV) virus free of heterologous elements,
PT and where protein gM has been deleted, useful as a vaccine for treating
PT or preventing EHV infections.
XX
PS Disclosure; SEQ ID NO 2; 156pp; English.
XX
CC The invention describes a new recombinant Equine Herpes Virus (EHV) where
CC the protein gM is absent, and the EHV is free of heterologous elements.
CC Also described are: a nucleic acid coding for an EHV defined above; a
CC vaccine preparation comprising the EHV or nucleic acid; obtaining a
CC recombinant EHV; and a cell line for use in the method, where the gene
CC encoding the protein gM is transfected into the cell line, and the cell
CC line expresses gM. The vaccine comprising the EHV or nucleic acid
CC encoding EHV is useful for treating and/or preventing EHV-associated
CC condition, and for monitoring the therapeutic success. The recombinant
CC EHV is useful as a vaccine against EHV infections. This sequence
CC represents the Equine herpesvirus 4 genome.
XX
SQ Sequence 145596 BP; 36213 A; 36932 C; 36529 G; 35922 T; 0 U; 0 Other;

Query Match 5.4%; Score 37.8; DB 12; Length 145596;
Best Local Similarity 61.9%; Pred. No. 4;
Matches 60; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

Qy 230 TACGCCGAACGGCCATGTTTGCCGAGACTGACCCTATGACATGGTTAAGACCAACGGTTG 289
| || ||| | |||||||||| | | ||| | || ||| | |||| | |
Db 46820 TTCGACGATCTGCCATGTTTGCGCACAACGACGCATCTACCTGGGCCCGGCCAAATATAG 46761

Qy 290 GCTTAAAACGTACGTTTAACCCGCGTATTATACGACC 326
| ||||||| || || ||||| |||| | ||
Db 46760 GACTAAAACGCACCTTCCCGCCGCGGTTTATGCAGCC 46724

RESULT 11
AAH24065
ID AAH24065 standard; DNA; 4590 BP.
XX
AC AAH24065;
XX
DT 29-AUG-2001 (first entry)
XX
DE Yeast AOD9604-associated DNA sequence, SEQ ID NO:1.
XX
KW Human growth hormone analogue peptide; hGH; AOD9604; lipid metabolism;
KW modulation; lipolysis stimulation; hormone-sensitive lipase stimulation;
KW lipogenesis inhibition; acetyl CoA carboxylase inhibition; obesity;
KW functional food; transgenic yeast; fat/lean ratio; food use; ds.
XX
OS Saccharomyces cerevisiae.
XX
FH Key Location/Qualifiers
FT misc_feature 10
FT /*tag= a
FT /note= "Represented as * in the specification"
FT misc_feature 3617
FT /*tag= b
FT /note= "Represented as * in the specification"
FT misc_feature 3649

```
FT                   /*tag=  c
FT                   /note= "Represented as * in the specification"
FT   misc_feature    3679
FT                   /*tag=  d
FT                   /note= "Represented as * in the specification"
FT   misc_feature    3819
FT                   /*tag=  e
FT                   /note= "Represented as * in the specification"
FT   misc_feature    3862
FT                   /*tag=  f
FT                   /note= "Represented as * in the specification"
FT   misc_feature    3864
FT                   /*tag=  g
FT                   /note= "Represented as * in the specification"
FT   misc_feature    3888
FT                   /*tag=  h
FT                   /note= "Represented as * in the specification"
FT   misc_feature    3890
FT                   /*tag=  i
FT                   /note= "Represented as * in the specification"
FT   misc_feature    3912
FT                   /*tag=  j
FT                   /note= "Represented as * in the specification"
FT   misc_feature    3914
FT                   /*tag=  k
FT                   /note= "Represented as * in the specification"
FT   misc_feature    3938
FT                   /*tag=  l
FT                   /note= "Represented as * in the specification"
FT   misc_feature    3939
FT                   /*tag=  m
FT                   /note= "Represented as * in the specification"
FT   misc_feature    3941
FT                   /*tag=  o
FT                   /note= "Represented as * in the specification"
FT   misc_feature    3943
FT                   /*tag=  p
FT                   /note= "Represented as * in the specification"
FT   misc_feature    4361
FT                   /*tag=  q
FT                   /note= "Represented as * in the specification"
XX
PN   WO200133977-A1.
XX
PD   17-MAY-2001.
XX
PF   06-NOV-2000; 2000WO-AU001362.
XX
PR   05-NOV-1999;   99AU-00003875.
XX
PA   (META-) METABOLIC PHARM LTD.
XX
PI   Belyea CI,  Ng FM,  Vaughan P;
XX
DR   WPI; 2001-328876/34.
XX
PT   New organisms containing nucleic acid encoding a growth hormone fragment
PT   which modulates lipid metabolism are useful to produce dietary aids for
PT   obesity and in the meat production industry.
XX
PS   Disclosure; Page 48-50; 54pp; English.
XX
CC   The invention relates to novel transgenic organisms useful in the
CC   production of functional food and drink products for the treatment or
CC   prevention of obesity via the regulation of lipid metabolism. The
CC   organisms comprise a polynucleotide encoding a growth hormone fragment
CC   capable of stimulating the activity of hormone-sensitive lipase (the key
CC   enzyme in lipolysis) and inhibiting acetyl CoA carboxylase (the key
CC   enzyme in lipogenesis). The growth hormone fragment preferably contains
CC   at least the disulphide-bonded loop of a mammalian growth hormone (but is
CC   not the full-length growth hormone) and is optionally linked to an
CC   epitope tag or heterologous fusion protein partner. The transgenic
CC   organism may be a microorganism used to produce a fermented product
CC   (e.g., yeast), or an edible plant or animal or cell thereof. Food or
CC   drink made using methods of the invention are used to modify fat/lean
CC   ratio, lipid metabolism or food use in a mammal. In particular, the food
CC   or drink products may be used to treat or prevent obesity, particularly
CC   in humans, and may also be used to improve the fat/lean ration of
CC   livestock raised for meat production. In the exemplification of the
CC   invention, the human growth hormone (hGH) fragment analogue AOD9604 was
CC   expressed in yeast, optionally fused to the FLAG epitope (AAB73625). The
CC   present sequence is described as a DNA sequence from yeast in the
CC   sequence listing, but is not further referred to in the specification
XX
SQ   Sequence 4590 BP; 661 A; 384 C; 127 G; 522 T; 0 U; 2896 Other;

  Query Match             5.3%;  Score 37.6;  DB 5;  Length 4590;
  Best Local Similarity   8.8%;  Pred. No. 1;
  Matches   33;  Conservative  181;  Mismatches  161;  Indels    0;  Gaps    0;

Qy          1 ATGACACAACCCGCATCGTCTCGTGTAGTCTTTGATCCCAGCAACCCCACCACATTTTCG 60
              :  :::: :::|::: ::::: ::: ::::| :::     :: ::::::   : : :
Db       4052 RNNMBRSCDMMCKMMMMKWYKNDBDVWRYYTGBBDRKNWNDYSMBSHMWNDYSKDCWNDY 4111

Qy         61 GTGGAAGCAATTGCGGCTTACACCCCCGTTGCTTTAATACGACTTTTAAACGCCAGTGGA 120
              :   | :|:::  : : : :: | :  :  ::   | :: : : |   : ::::|: :::
Db       4112 SCTMATRCVRYSAVASDWNDYSCTMATRCVRYSAVASDWNDYSCTMATRCVRYSAVASDW 4171

Qy        121 CCTTTGCAACCTGGTCACCGTGTGGACATCGCTGATGCCAGAAGCATTTACACCGTGGGA 180
                :   :|  | :   : ::: : :  :  :| :   :  :: ::  | | :|:::: :|
Db       4172 NDYSCTMATRCVRYSAVASDWNDYSCTMATRCVRYSAVASDWNDYSCTMATRCVRYSAVA 4231

Qy        181 GCCGCGGCCAGTGCCGCGCGTGCACGCGCTAACCATAATGCAAATACGATACGCCGAACG 240
              :  :::  : :::::|  : :::| :: : :|| : ::::: ::  :: :: : |: :
Db       4232 SDRKSKWWSKKWSSMGTMSNWRMANSYMBARACRRNWHHBBSDHMBSSTDWRDDCSTMAT 4291

Qy        241 GCCATGTTTGCCGAGACTGACCCTATGACATGGTTAAGACCAACGGTTGGCTTAAAACGT 300
              ::|:|  :::  : ::: ::  :  : :: : :: :: :: :: :: :|: ::|   : |
Db       4292 SSCDTWNDYSWNDYSHHSSWNDYSNDNDSCWNDYSDMCRSTWRDRRSDGRTHHACKCHTT 4351

Qy        301 ACGTTTAACCCGCGTATTATACGACCACAACCCCCAAATCCATCCATGAGTTTGGGAATC 360
              : :  :::     ::: :| :::     |    : ::   :    : :::|: :: : ||
Db       4352 DGDANBDDDRTNTRYRSHARRSRBTABRACATNSADMTADMNSTDWRDDSTWRDDCMNTC 4411

Qy        361 TCGGGGCCTACTATA 375
                ::   |:|: |:
Db       4412 MNTSRCDNTMCDNTM 4426


RESULT 12
ACA43990/c
ID   ACA43990 standard; DNA; 711 BP.
XX
AC   ACA43990;
XX
DT   19-JUN-2003  (first entry)
XX
DE   Prokaryotic essential gene #25647.
XX
KW   Antisense; ds; prokaryotic essential gene; cell proliferation;
KW   drug design; gene.
XX
OS   Pseudomonas putida.
XX
PN   WO200277183-A2.
XX
PD   03-OCT-2002.
XX
PF   21-MAR-2002; 2002WO-US009107.
XX
PR   21-MAR-2001; 2001US-00815242.
PR   06-SEP-2001; 2001US-00948993.
PR   25-OCT-2001; 2001US-0342923P.
PR   08-FEB-2002; 2002US-00072851.
PR   06-MAR-2002; 2002US-0362699P.
XX
```

PA (ELIT-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
DR WPI; 2003-029926/02.
DR P-PSDB; ABU40120.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 14; SEQ ID NO 31860; 1766pp; English.
XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide;(5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than S. aureus, S. typhimurium,
CC K. pneumoniae or P. aeruginosa. The present sequence is one of the target
CC prokaryotic essential genes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 711 BP; 120 A; 161 C; 304 G; 126 T; 0 U; 0 Other;

Query Match 5.2%; Score 37; DB 8; Length 711;
Best Local Similarity 47.3%; Pred. No. 0.72;
Matches 112; Conservative 0; Mismatches 125; Indels 0; Gaps 0;

```
Qy      428 CGTTGGCGTTTTCGGGATCATCCCCGCAACACCCCCCACCTCAAACGACGTCGGCATCCG 487
            || |  ||    |  |  || | |||| ||  |  ||||| | | ||  | |  || |
Db      400 CGCTTCCGCCACCATGGCCACCGCCGCCACCACTTCCACCACCACCGTGGCCACCACCAC 341

Qy      488 TTGGACAACAGCAACACGTGGTGTCGGGGTCTTCTGGACAACAACCGCAACAGGGAGCAC 547
                | | |||   | | |   | | |   | | |||| | || | || |||
Db      340 CGCCGCTTCCACCACCACCGTGGCCACCGCCACCGCCGCTTCCACCACCACCGTGACCAC 281

Qy      548 AGTCAAGCACTGTCCAGCCAACAACCGGATCACCGCCCGCGGCCCAAGGCGTGCCACAGT 607
             | ||    |  | |  ||  || |  | ||||||  || | |     |||| ||
Db      280 CGCCACCGCCGCTTCCACCGCCACCGTGGCCACCGCCGCCGCCGCTTCCACCGCCACCGT 221

Qy      608 CTACCCCGCCCCCAACCCAAAATACCCCCCAGGGGGGTAAGGGACAGACCTTGTCAC 664
              | |||||  ||  | |      | || | | ||     | || | |   | |||
Db      220 GGCCACCGCCGCCGCCACTTCCACCGCCACCGTGGCCACCGCCACCGCCACCGCCAC 164
```

RESULT 13
AAI99682_27
Continuation (28 of 45) of AAI99682 from base 2700001 (Mycobacterium tuberculosis strain
WP Sequence split into 45 fragments LOCUS AAI99682 Accession Aai99682

WP	Fragment Name	Begin	End
WP	AAI99682_00	1	110000
WP	AAI99682_01	100001	210000
WP	AAI99682_02	200001	310000
WP	AAI99682_03	300001	410000
WP	AAI99682_04	400001	510000
WP	AAI99682_05	500001	610000
WP	AAI99682_06	600001	710000
WP	AAI99682_07	700001	810000
WP	AAI99682_08	800001	910000
WP	AAI99682_09	900001	1010000
WP	AAI99682_10	1000001	1110000
WP	AAI99682_11	1100001	1210000
WP	AAI99682_12	1200001	1310000
WP	AAI99682_13	1300001	1410000
WP	AAI99682_14	1400001	1510000
WP	AAI99682_15	1500001	1610000
WP	AAI99682_16	1600001	1710000
WP	AAI99682_17	1700001	1810000
WP	AAI99682_18	1800001	1910000
WP	AAI99682_19	1900001	2010000
WP	AAI99682_20	2000001	2110000
WP	AAI99682_21	2100001	2210000
WP	AAI99682_22	2200001	2310000
WP	AAI99682_23	2300001	2410000
WP	AAI99682_24	2400001	2510000
WP	AAI99682_25	2500001	2610000
WP	AAI99682_26	2600001	2710000
WP	AAI99682_27	2700001	2810000
WP	AAI99682_28	2800001	2910000
WP	AAI99682_29	2900001	3010000
WP	AAI99682_30	3000001	3110000
WP	AAI99682_31	3100001	3210000
WP	AAI99682_32	3200001	3310000
WP	AAI99682_33	3300001	3410000
WP	AAI99682_34	3400001	3510000
WP	AAI99682_35	3500001	3610000
WP	AAI99682_36	3600001	3710000
WP	AAI99682_37	3700001	3810000
WP	AAI99682_38	3800001	3910000
WP	AAI99682_39	3900001	4010000
WP	AAI99682_40	4000001	4110000
WP	AAI99682_41	4100001	4210000
WP	AAI99682_42	4200001	4310000
WP	AAI99682_43	4300001	4410000
WP	AAI99682_44	4400001	4411529

Query Match 5.1%; Score 36.2; DB 4; Length 110000;
Best Local Similarity 45.6%; Pred. No. 11;
Matches 128; Conservative 0; Mismatches 153; Indels 0; Gaps 0;

```
Qy      397 GCCGATCAGTCTGCTTTACAACAGCCCGCCGCGTTGGCGTTTTCGGGATCATCCCCGCAA 456
            ||||| | | | ||    |  |||  || ||  | ||||| ||    || | ||
Db   105126 GCCGACCCGCCAGCACCCCCGGCGCCGCCCACGCCGCCGTTTCCGCCGGCACCGCCATTG 105185

Qy      457 CACCCCCCACCTCAAACGACGTCGGCATCCGTTGGACAACAGCAACACGTGGTGTCGGGG 516
            |  ||  | || |   |  || ||||  |||    ||       |   |  |  |||
Db   105186 CCGCCGGCGCCGCCGTCACCGCCGGCCGCCGAGATACCGGCCGGCCCATTTATTCCGGTA 105245

Qy      517 TCTTCTGGACAACAACCGCAACAGGGAGCACAGTCAAGCACTGTCCAGCCAACAACCGGA 576
            | | | ||  |   | |  | ||  |||| | ||    | | || ||| ||| |   |
Db   105246 GCCCCGGCACCGCCGGCACCGCCGGCCGCACCGGCACCACCGGCCCCGCCGACACCGCCA 105305

Qy      577 TCACCGCCCGCGGCCCAAGGCGTGCCACAGTCTACCCCGCCCCCAACCCAAAATACCCCC 636
             | || || ||| | |     |  |     | ||  |  |||  |         || ||
Db   105306 ACGCCACCGGCGCCGCCGTTACCGAGAAGCCACCCTCCCCGACCACCGTTGCCGCCCACC 105365

Qy      637 CAGGGGGGTAAGGGACAGACCTTGTCACACACGGGACAATC 677
            |    ||    | |  | ||  ||| || | |   ||| |
Db   105366 CCACCGGCACCGCCATCGCCCCCGTCCGACCCTGCCAAACC 105406
```

RESULT 14

AAI99682_28
Continuation (29 of 45) of AAI99682 from base 2800001 (Mycobacterium tuberculosis strain
WP Sequence split into 45 fragments LOCUS AAI99682 Accession Aai99682

WP	Fragment Name	Begin	End
WP	AAI99682_00	1	110000
WP	AAI99682_01	100001	210000
WP	AAI99682_02	200001	310000
WP	AAI99682_03	300001	410000
WP	AAI99682_04	400001	510000
WP	AAI99682_05	500001	610000
WP	AAI99682_06	600001	710000
WP	AAI99682_07	700001	810000
WP	AAI99682_08	800001	910000
WP	AAI99682_09	900001	1010000
WP	AAI99682_10	1000001	1110000
WP	AAI99682_11	1100001	1210000
WP	AAI99682_12	1200001	1310000
WP	AAI99682_13	1300001	1410000
WP	AAI99682_14	1400001	1510000
WP	AAI99682_15	1500001	1610000
WP	AAI99682_16	1600001	1710000
WP	AAI99682_17	1700001	1810000
WP	AAI99682_18	1800001	1910000
WP	AAI99682_19	1900001	2010000
WP	AAI99682_20	2000001	2110000
WP	AAI99682_21	2100001	2210000
WP	AAI99682_22	2200001	2310000
WP	AAI99682_23	2300001	2410000
WP	AAI99682_24	2400001	2510000
WP	AAI99682_25	2500001	2610000
WP	AAI99682_26	2600001	2710000
WP	AAI99682_27	2700001	2810000
WP	AAI99682_28	2800001	2910000
WP	AAI99682_29	2900001	3010000
WP	AAI99682_30	3000001	3110000
WP	AAI99682_31	3100001	3210000
WP	AAI99682_32	3200001	3310000
WP	AAI99682_33	3300001	3410000
WP	AAI99682_34	3400001	3510000
WP	AAI99682_35	3500001	3610000
WP	AAI99682_36	3600001	3710000
WP	AAI99682_37	3700001	3810000
WP	AAI99682_38	3800001	3910000
WP	AAI99682_39	3900001	4010000
WP	AAI99682_40	4000001	4110000
WP	AAI99682_41	4100001	4210000
WP	AAI99682_42	4200001	4310000
WP	AAI99682_43	4300001	4410000
WP	AAI99682_44	4400001	4411529

Query Match 5.1%; Score 36.2; DB 4; Length 110000;
Best Local Similarity 45.6%; Pred. No. 11;
Matches 128; Conservative 0; Mismatches 153; Indels 0; Gaps 0;

```
Qy        397 GCCGATCAGTCTGCTTTACAACAGCCCGCCGCGTTGGCGTTTTCGGGATCATCCCCGCAA 456
              ||||| | | | ||    |   |||  || ||  | ||||| ||    || | ||
Db       5126 GCCGACCCGCCAGCACCCCCGGCGCCGCCCACGCCGCCGTTTCCGCCGGCACCGCCATTG 5185

Qy        457 CACCCCCCCACCTCAAACGACGTCGGCATCCGTTGGACAACAGCAACACGTGGTGTCGGGG 516
              |  ||  | || |   |  || ||||  |||    ||         |   |  | |||
Db       5186 CCGCCGGCGCCGCCGTCACCGCCGGCCGCCGAGATACCGGCCGGCCCATTTATTCCGGTA 5245

Qy        517 TCTTCTGGACAACAACCGCAACAGGGAGCACAGTCAAGCACTGTCCAGCCAACAACCGGA 576
              | | | ||  |   | | | ||  |||| | ||   | | || ||| ||| |   |
Db       5246 GCCCCGGCACCGCCGGCACCGCCGGCCGCACCGGCACCACCGGCCCCGCCGACACCGCCA 5305

Qy        577 TCACCGCCCGCGGCCCAAGGCGTGCCACAGTCTACCCCGCCCCCAACCCAAAATACCCCC 636
              | || || ||| | |       | |       | || |  ||| |         || ||
Db       5306 ACGCCACCGGCGCCGCCGTTACCGAGAAGCCACCCTCCCCGACCACCGTTGCCGCCCACC 5365

Qy        637 CAGGGGGGTAAGGGACAGACCTTGTCACACACGGGACAATC 677
              |    ||    | | | ||  |||  || | |   || |
Db       5366 CCACCGGCACCGCCATCGCCCCCGTCCGACCCTGCCAAACC 5406
```

RESULT 15
AAI99683_27
Continuation (28 of 44) of AAI99683 from base 2700001 (Mycobacterium tuberculosis strain
WP Sequence split into 44 fragments LOCUS AAI99683 Accession Aai99683

WP	Fragment Name	Begin	End
WP	AAI99683_00	1	110000
WP	AAI99683_01	100001	210000
WP	AAI99683_02	200001	310000
WP	AAI99683_03	300001	410000
WP	AAI99683_04	400001	510000
WP	AAI99683_05	500001	610000
WP	AAI99683_06	600001	710000
WP	AAI99683_07	700001	810000
WP	AAI99683_08	800001	910000
WP	AAI99683_09	900001	1010000
WP	AAI99683_10	1000001	1110000
WP	AAI99683_11	1100001	1210000
WP	AAI99683_12	1200001	1310000
WP	AAI99683_13	1300001	1410000
WP	AAI99683_14	1400001	1510000
WP	AAI99683_15	1500001	1610000
WP	AAI99683_16	1600001	1710000
WP	AAI99683_17	1700001	1810000
WP	AAI99683_18	1800001	1910000
WP	AAI99683_19	1900001	2010000
WP	AAI99683_20	2000001	2110000
WP	AAI99683_21	2100001	2210000
WP	AAI99683_22	2200001	2310000
WP	AAI99683_23	2300001	2410000
WP	AAI99683_24	2400001	2510000
WP	AAI99683_25	2500001	2610000
WP	AAI99683_26	2600001	2710000
WP	AAI99683_27	2700001	2810000
WP	AAI99683_28	2800001	2910000
WP	AAI99683_29	2900001	3010000
WP	AAI99683_30	3000001	3110000
WP	AAI99683_31	3100001	3210000
WP	AAI99683_32	3200001	3310000
WP	AAI99683_33	3300001	3410000
WP	AAI99683_34	3400001	3510000
WP	AAI99683_35	3500001	3610000
WP	AAI99683_36	3600001	3710000
WP	AAI99683_37	3700001	3810000
WP	AAI99683_38	3800001	3910000
WP	AAI99683_39	3900001	4010000
WP	AAI99683_40	4000001	4110000
WP	AAI99683_41	4100001	4210000
WP	AAI99683_42	4200001	4310000
WP	AAI99683_43	4300001	4403765

Query Match 5.1%; Score 36.2; DB 4; Length 110000;
Best Local Similarity 45.6%; Pred. No. 11;
Matches 128; Conservative 0; Mismatches 153; Indels 0; Gaps 0;

```
Qy        397 GCCGATCAGTCTGCTTTACAACAGCCCGCCGCGTTGGCGTTTTCGGGATCATCCCCGCAA 456
              ||||| | | | ||    |   |||  || ||  | ||||| ||    || | ||
Db     100579 GCCGACCCGCCAGCACCCCCGGCGCCGCCCACGCCGCCGTTTCCGCCGGCACCGCCATTG 100638

Qy        457 CACCCCCCCACCTCAAACGACGTCGGCATCCGTTGGACAACAGCAACACGTGGTGTCGGGG 516
              |  ||  | || |   |  || ||||  |||    ||         |   |  | |||
Db     100639 CCGCCGGCGCCGCCGTCACCGCCGGCCGCCGAGATACCGGCCGGCCCATTTATTCCGGTA 100698

Qy        517 TCTTCTGGACAACAACCGCAACAGGGAGCACAGTCAAGCACTGTCCAGCCAACAACCGGA 576
              | | | ||  |   | | | ||  |||| | ||   | | || ||| ||| |   |
Db     100699 GCCCCGGCACCGCCGGCACCGCCGGCCGCACCGGCACCACCGGCCCCGCCGACACCGCCA 100758

Qy        577 TCACCGCCCGCGGCCCAAGGCGTGCCACAGTCTACCCCGCCCCCAACCCAAAATACCCCC 636
              | || || ||| | |       | |       | || |  ||| |         || ||
Db     100759 ACGCCACCGGCGCCGCCGTTACCGAGAAGCCACCCTCCCCGACCACCGTTGCCGCCCACC 100818

Qy        637 CAGGGGGGTAAGGGACAGACCTTGTCACACACGGGACAATC 677
```

```
                |    ||    |  |  | ||  |||  || | |   || |
Db     100819 CCACCGGCACCGCCATCGCCCCCGTCCGACCCTGCCAAACC 100859
```

Search completed: September 25, 2005, 21:51:41
Job time : 538 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: September 25, 2005, 21:30:54 ; Search time 3219 Seconds
(without alignments)
8336.536 Million cell updates/sec

Title: US-09-874-140-1
Perfect score: 705
Sequence: 1 atgacacaacccgcatcgtc.....cttcaagaagtcgtagggtg 705

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_htc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	45.6	6.5	434	9	CNS049AP	AL280330 Tetraodon
c	2	44.8	6.4	884	9	CNS006U0	AL065923 Drosophil
	3	44.4	6.3	1013	9	CNS016KT	AL106871 Drosophil
c	4	42.8	6.1	970	9	CNS010C9	AL098787 Drosophil
c	5	42.2	6.0	977	9	CNS00JX7	AL076850 Drosophil
	6	41.8	5.9	1235	4	BM473847	BM473847 AGENCOURT
	7	40.8	5.8	710	9	AG135197	AG135197 Pan trogl
	8	40.2	5.7	1156	9	AG075888	AG075888 Pan trogl
	9	40.2	5.7	1201	9	CNS0169V	AL106477 Drosophil
c	10	40	5.7	910	9	CNS006ON	AL065629 Drosophil
	11	39.6	5.6	454	9	CE497548	CE497548 tigr-gss-
	12	39.6	5.6	1055	8	BZ144309	CH230-462
c	13	39.4	5.6	870	9	AG471975	AG471975 Mus muscu
c	14	39.4	5.6	922	9	CNS0073W	AL066784 Drosophil
	15	39.2	5.6	787	9	CC910437	CC910437 t061h01ba
c	16	39.2	5.6	942	6	CB314169	CB314169 AGENCOURT
	17	39.2	5.6	1537	5	BU372422	BU372422 603811129
c	18	39	5.5	696	8	BZ365757	BZ365757 ie28h05.g
c	19	39	5.5	970	1	AL519510	AL519510 AL519510
	20	38.4	5.4	961	9	CNS01087	AL098641 Drosophil
	21	38.4	5.4	1606	4	BG027840	BG027840 602294810
c	22	38.2	5.4	600	8	BZ895383	BZ895383 NaRP10_01
	23	38.2	5.4	1101	9	CNS00KK2	AL077673 Drosophil
c	24	38	5.4	529	9	CNS00LHZ	AL068241 Drosophil
c	25	38	5.4	602	1	AI727059	AI727059 BNLGHi724
c	26	38	5.4	752	7	CO096155	CO096155 GR__Ea19D
c	27	38	5.4	804	7	CO107101	CO107101 GR__Eb003
c	28	38	5.4	806	9	AG081452	AG081452 Pan trogl
c	29	38	5.4	808	7	CO118443	CO118443 GR__Eb020
c	30	38	5.4	815	7	CO123895	CO123895 GR__Eb06G
	31	37.8	5.4	392	1	AI397669	AI397669 NCSC5H5T3
	32	37.8	5.4	623	7	CF421899	CF421899 NcEST3d87
c	33	37.8	5.4	841	9	AG388909	AG388909 Mus muscu
	34	37.8	5.4	869	5	BU959836	BU959836 AGENCOURT
	35	37.8	5.4	878	4	BG034409	BG034409 602302810
	36	37.8	5.4	939	9	CNS00CNG	AL059400 Drosophil
	37	37.6	5.3	936	9	CNS01608	AL106130 Drosophil
	38	37.4	5.3	636	7	CO080152	CO080152 GR__Ea43G
	39	37.4	5.3	936	5	BU151687	BU151687 AGENCOURT
	40	37.4	5.3	1097	7	CK212140	CK212140 FGAS02400
	41	37.4	5.3	1473	9	CL972108	CL972108 OsIFCC041
	42	37.2	5.3	410	1	AU298532	AU298532 AU298532
c	43	37.2	5.3	909	9	CNS00JTL	AL076720 Drosophil
	44	37.2	5.3	953	9	CNS017Y4	AL108646 Drosophil
	45	37.2	5.3	964	9	CNS003WG	AL065254 Drosophil

ALIGNMENTS

RESULT 1
CNS049AP
LOCUS CNS049AP 434 bp DNA linear GSS 01-SEP-2000
DEFINITION Tetraodon nigroviridis genome survey sequence PUC-Ori end of clone 093G05 of library G from Tetraodon nigroviridis, genomic survey sequence.
ACCESSION AL280330
VERSION AL280330.1 GI:8018616
KEYWORDS GSS; genome survey sequence.
SOURCE Tetraodon nigroviridis
ORGANISM Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetradontoidea; Tetraodontidae; Tetraodon.
REFERENCE 1
AUTHORS Roest Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C., Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F., Saurin,W. and Weissenbach,J.
TITLE Estimate of human gene number provided by genome-wide analysis using Tetraodon nigroviridis DNA sequence
JOURNAL Nat. Genet. 25 (2), 235-238 (2000)
MEDLINE 20296633
PUBMED 10835645
REFERENCE 2
AUTHORS Roest Crollius,H., Jaillon,O., Dasilva,C., Ozouf-Costaz,C., Fizames,C., Fischer,C., Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and Weissenbach,J.
TITLE Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis
JOURNAL Genome Res. 10 (7), 939-949 (2000)
MEDLINE 20359837
PUBMED 10899143
REFERENCE 3 (bases 1 to 434)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (12-APR-2000) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at http://www.genoscope.cns.fr/Tetraodon.
FEATURES Location/Qualifiers
source 1. .434
/organism="Tetraodon nigroviridis"
/mol_type="genomic DNA"

```
                     /db_xref="taxon:99883"
                     /clone="093G05"
                     /clone_lib="G"
                     /note="Genoscope sequence ID : C0BG093AD03SP1~end :
                     PUC-Ori"
ORIGIN

  Query Match              6.5%;  Score 45.6;  DB 9;  Length 434;
  Best Local Similarity   41.9%;  Pred. No. 0.03;
  Matches   96;  Conservative   24;  Mismatches  109;  Indels    0;  Gaps    0;

Qy        409 GCTTTACAACAGCCCGCCGCGTTGGCGTTTTCGGGATCATCCCCGCAACACCCCCCACCT 468
              |||  |    | |   :  ||| | |||   ||  ||:| : ||| | || ||| || ||
Db         72 GCTGGAGCGGACCGGNMTCCGTCGTCGTCSCCGCCATMASMGCCGGAGCAGCCCTCAGCT 131

Qy        469 CAAACGACGTCGGCATCCGTTGGACAACAGCAACACGTGGTGTCGGGGTCTTCTGGACAA 528
              | | || ||||  | |    :  |  :|| | | |  ||   |   |:| :|    | |
Db        132 CCATCGTCGTCCCCGCCATCASMGCCGSAGAAGCCCYCGGCTCCMTCGYCSYCACCGCCA 191

Qy        529 CAACCGCAACAGGGAGCACAGTCAAGCACTGTCCAGCCAACAACCGGATCACCGCCCGCG 588
              : |: ||  :||    :  :  |      ||  |: : |  |  ||  | ||| |||   |
Db        192 YCASAGCCGSAGCCGCMGAMCCCCCCGGCTCCCYCSTCCCCGCCCCCAACACAGCCGCAG 251

Qy        589 GCCCAAGGCGTGCCACAGTCTACCCCGCCCCCAACCCAAAATACCCCCC 637
              ::|        |||: ::|  | ||:||:||  :||   : :||||||
Db        252 AMSCCCCCGCGTCCAYCSYCCCCGCCSCCMCCCCMCCCCCMCMCCCCCC 300


RESULT 2
CNS006U0/c
LOCUS       CNS006U0                  884 bp    DNA     linear   GSS 03-JUN-1999
DEFINITION  Drosophila melanogaster genome survey sequence T7 end of BAC #
            BACR14N21 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL065923
VERSION     AL065923.1  GI:4944891
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster (fruit fly)
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 884)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see  http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutoyo Osoegawa and
            Aaron Mammoser in Pieter de Jong's laboratory in the Department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RPCI-98 and was constructed by partial
            EcoRI digestion of Drosophila DNA provided by the BDGP from the
            isogenic strain y2; cn bw sp, the same strain used for the BDGP's
            P1 and EST libraries. A more detailed description of the library
            and how to order individual BAC clones, the entire library, or
            filters for hybridization from the BACPAC Resource Center can be
            found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES             Location/Qualifiers
     source          1. .884
                     /organism="Drosophila melanogaster"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7227"
                     /clone="BACR14N21"
                     /clone_lib="RPCI-98"
                     /note="end : T7"
ORIGIN

  Query Match              6.4%;  Score 44.8;  DB 9;  Length 884;
  Best Local Similarity   18.7%;  Pred. No. 0.062;
  Matches   75;  Conservative  141;  Mismatches  185;  Indels    0;  Gaps    0;

Qy        294 AAAACGTACGTTTAACCCGCGTATTATACGACCACAACCCCCCAAATCCATCCATGAGTTT 353
              ::::: :  ::::  :::::  :  : :: ::::::::::::::  ::: :::  :
Db        880 MMVMMMHTKKKKKTTHMMVMNMMMMMMMMMMMMMMMMMMMMMMGMMMMMMMMMMAMMM 821

Qy        354 GGGAATCTCGGGGGCCTACTATATTGCCGCAAAAAACACAGAGCGCCGATCAGTCTGCTTT 413
              :: : :    :: :: : :   :: ::::|::|::: : : || : :|  |  :
Db        820 MMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMAMMCMMMCMMMACCMMCMAMMCAMMMMM 761

Qy        414 ACAACAGCCCGCCGCGTTGGCGTTTTCGGGATCATCCCCGCAACACCCCCCACCTCAAAC 473
              ::::: ::: :: :      :      |   : |: :: | |:|||||:| ||:::    ::
Db        760 CMMMMMMMMCMMMMMCMMCMCCMMMCCMCMCCMMMMACACMACACMCACCMMMMACSMM 701

Qy        474 GACGTCGGCATCCGTTGGACAACAGCAACACGTGGTGTCGGGGTCTTCTGGACAACAACC 533
              |:  |  :| ::      : |::: || ::|        :      |  |   :|:::: ||
Db        700 MAMCMCMCMACMMMMMMMMAAMMMCCACHMCCAMACCMMMCCCCCCMCACCMCMMMMCCC 641

Qy        534 GCAACAGGGAGCACAGTCAAGCACTGTCCAGCCAACAACCGGATCACCGCCCGCGGCCCA 593
              |  |:   |  |||   |: :|:    :| : :| |:||  : |  | | | :  :|||
Db        640 CCCCCMMMMACAACAMCAAMAMAMACMAMAAAMAMAAAMCCMCMCCCACCCACCMCCMCCA 581

Qy        594 AGGCGTGCCACAGTCTACCCCGCCCCCAACCCAAAATACCCCCCAGGGGGGTAAGGGACA 653
              :  |      ::   :::   :  :||   ::|: :| :||:: ||          | :
Db        580 MCTCACCAAMMCAAHHMAAAMMAMCCABCMMCHCMASMCCMMACAYMCCCCCCAMBTBKC 521

Qy        654 GACCTTGTCACACACGGGACAATCTGGAAACGCTTCAAGAA 694
              ::|:::  | | ||:   |:||:|: |   | |   || ||
Db        520 BMCMYBCCCCCCCAMMAMAMAAKCKCGGGSCACCMAAAAAA 480


RESULT 3
CNS016KT
LOCUS       CNS016KT                  1013 bp    DNA     linear   GSS 26-JUL-1999
DEFINITION  Drosophila melanogaster genome survey sequence SP6 end of BAC
            BACN16J16 of DrosBAC library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL106871
VERSION     AL106871.1  GI:5624218
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster (fruit fly)
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 1013)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the European Drosophila Genome Project (EDGP) -
            http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
            library (Dros BAC) was made by Alain Billaud at CEPH (Centre
            d'Etude du Polymorphisme Humain) with funding provided by a MRC
            project grant. The DNA was prepared from embryos by Alain Bucheton
            and Genevieve Payan. It has been constructed in the vector
            pBeloBAC11.
FEATURES             Location/Qualifiers
     source          1. .1013
                     /organism="Drosophila melanogaster"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7227"
                     /clone="BACN16J16"
                     /clone_lib="DrosBAC"
                     /plasmid="pBeloBAC11"
                     /note="end : SP6"
ORIGIN

  Query Match              6.3%;  Score 44.4;  DB 9;  Length 1013;
```

```
  Best Local Similarity   15.0%;  Pred. No. 0.084;
  Matches   37;  Conservative   93;  Mismatches  116;  Indels    0;  Gaps    0;

Qy        415 CAACAGCCCGCCGCGTTGGCGTTTTCGGGATCATCCCCGCAACACCCCCCACCTCAAACG 474
              |: :::||| :  : :  :|     :| : :  :|:| : :::::::::: : | : :
Db        246 CMTMMSCCCNMNCMMYSCSCCCCMSMGMMMCMNMMCMCCMNMMMMSMMMVMNMNCGMGMT 305

Qy        475 ACGTCGGCATCCGTTGGACAACAGCAACACGTGGTGTCGGGGTCTTCTGGACAACAACCG 534
              |:   |     : |  ::::     : :|    |        |  :   :::::::::
Db        306 AMCMTGNNTNGMMGTTHMMMMGNTTTMNGMGGNMNGNGMNNNNCNGMNNNMMMMMMMMM 365

Qy        535 CAACAGGGAGCACAGTCAAGCACTGTCCAGCCAACAACCGGATCACCGCCCGCGGCCCAA 594
               : |:|||:|:: : ||  |::    : :  :: |: ::  : :::: ::: :  |:|::
Db        366 GMGCMGGGMGMMGMNTCTTGMMGMCNMGMCNMMNCMCMMMMMMMMMMMMMMCMMMCMCMM 425

Qy        595 GGCGTGCCACAGTCTACCCCGCCCCCAACCCAAAATACCCCCCAGGGGGGTAAGGGACAG 654
                :   :: :   : :::|: ::::::      :   :|: :  || |    | |  :|
Db        426 TMMMMMMMCMCMMMMMMMCMMMMMMMMCGGNNGGMNTGMCMNMGMGGNGNGGNGNGNNMG 485

Qy        655 ACCTTG 660
                 ||:|
Db        486 GTCTYG 491


RESULT 4
CNS010C9/c
LOCUS       CNS010C9                970 bp    DNA     linear   GSS 26-JUL-1999
DEFINITION  Drosophila melanogaster genome survey sequence SP6 end of BAC
            BACN03L05 of DrosBAC library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL098787
VERSION     AL098787.1  GI:5610398
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster (fruit fly)
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 970)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the European Drosophila Genome Project (EDGP) -
            http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
            library (Dros BAC) was made by Alain Billaud at CEPH (Centre
            d'Etude du Polymorphisme Humain) with funding provided by a MRC
            project grant. The DNA was prepared from embryos by Alain Bucheton
            and Genevieve Payan. It has been constructed in the vector
            pBeloBAC11.
FEATURES             Location/Qualifiers
     source          1. .970
                     /organism="Drosophila melanogaster"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7227"
                     /clone="BACN03L05"
                     /clone_lib="DrosBAC"
                     /plasmid="pBeloBAC11"
                     /note="end : SP6"
ORIGIN

  Query Match             6.1%;  Score 42.8;  DB 9;  Length 970;
  Best Local Similarity   23.1%;  Pred. No. 0.25;
  Matches   42;  Conservative   77;  Mismatches   63;  Indels    0;  Gaps    0;

Qy        463 CCACCTCAAACGACGTCGGCATCCGTTGGACAACAGCAACACGTGGTGTCGGGGTCTTCT 522
              :: :: : :|:::::  : :| ::   :::::|:|::|: ||     : :::::
Db        922 SMSVMCSSVASSRSSVARAVASVSCARRSMSRAVARVARGACASAASRSVRVVRVRARAG 863

Qy        523 GGACAACAACCGCAACAGGGAGCACAGTCAAGCACTGTCCAGCCAACAACCGGATCACCG 582
              ::|:|: :|:   ::    :|| | ||   :::|: |  :|:::: :  : |:: : ||:

Db        862 VRAMARRRASAARMVACAARAGAARAGARGRSVAVVGVRSAVSSVSSCSSAGRVMVCCCR 803

Qy        583 CCCGCGGCCCAAGGCGTGCCACAGTCTACCCCGCCCCCAACCCAAAATACCCCCCAGGGG 642
              | |:: :|:|  : |:| ||: ::  :  ||:| ||:||| :||: ::  :|:|: :: :
Db        802 CGCRSASCVCGMCGSGSGCSCMMCASCSCCMCCCCMCCAGSCCMCMMCGSCSCSGMSMVM 743

Qy        643 GG 644
              :|
Db        742 SG 741


RESULT 5
CNS00JX7/c
LOCUS       CNS00JX7                977 bp    DNA     linear   GSS 03-JUN-1999
DEFINITION  Drosophila melanogaster genome survey sequence TET3 end of BAC:
            BACR39C01 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL076850
VERSION     AL076850.1  GI:4956428
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster (fruit fly)
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 977)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see  http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutoyo Osoegawa and
            Aaron Mammoser in Pieter de Jong's laboratory in the Department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RPCI-98 and was constructed by partial
            EcoRI digestion of Drosophila DNA provided by the BDGP from the
            isogenic strain y2; cn bw sp, the same strain used for the BDGP's
            P1 and EST libraries. A more detailed description of the library
            and how to order individual BAC clones, the entire library, or
            filters for hybridization from the BACPAC Resource Center can be
            found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES             Location/Qualifiers
     source          1. .977
                     /organism="Drosophila melanogaster"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7227"
                     /clone="BACR39C01"
                     /clone_lib="RPCI-98"
                     /note="end : TET3"
ORIGIN

  Query Match             6.0%;  Score 42.2;  DB 9;  Length 977;
  Best Local Similarity   15.0%;  Pred. No. 0.38;
  Matches   40;  Conservative  106;  Mismatches  121;  Indels    0;  Gaps    0;

Qy        357 AATCTCGGGGCCTACTATATTGCCGCAAAAAACACAGAGCGCCGATCAGTCTGCTTTACA 416
              || | ::::::: ::  :| :    || |||:  :  |:    :||:|: :|   | |:|
Db        378 AAGCASRSSSSSCRSACYAAYAAATCATAAAWATBSCASAWAASATSAKCSTAAATAASA 319

Qy        417 ACAGCCCGCCGCGTTGGCGTTTTCGGGATCATCCCCGCAACACCCCCCACCTCAAACGAC 476
              :|:   ::::::  :::     ::::: :| ||:::   :: :::::| : : ||:: :
Db        318 SSABNAAVSVSSSAASSSACSASSSSSVSSASCCSSSASSSVASSSSSAASSSSAASSSS 259

Qy        477 GTCGGCATCCGTTGGACAACAGCAACACGTGGTGTCGGGGTCTTCTGGACAACAACCGCA 536
              : :::: |:::  ::  :|  ::  :    ::  ::::: :  : :: : | | :: :
Db        258 SSSSVSSTSSSSVSSSAVAATSSSSBSAAASSSAASSSSSSSSSSVSSSSSAAASSSASS 199

Qy        537 ACAGGGAGCACAGTCAAGCACTGTCCAGCCAACAACCGGATCACCGCCCGCGGCCCAAGG 596
```

```
Db        198 SSATSSASSAASSSSSAASSSAATSASSCSSSSTCSSSSSBTCSSSSSVAGANCNSNNGN 139

Qy        597 CGTGCCACAGTCTACCCCGCCCCCAAC 623
                  |               ||  |  :
Db        138 NNANCNNNNNNNNNNNNCCCNNCNNS 112


RESULT 6
BM473847
LOCUS       BM473847               1235 bp    mRNA    linear   EST 05-FEB-2002
DEFINITION  AGENCOURT_6467956 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:5582144
            5', mRNA sequence.
ACCESSION   BM473847
VERSION     BM473847.1  GI:18522889
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1235)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12343  row: c  column: 09
            High quality sequence stop: 217.
FEATURES             Location/Qualifiers
     source          1. .1235
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5582144"
                     /tissue_type="retinoblastoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_67"
                     /note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.75 kb.   Library constructed by Life
                     Technologies."
ORIGIN

  Query Match             5.9%;  Score 41.8;  DB 4;  Length 1235;
  Best Local Similarity   47.8%;  Pred. No. 0.52;
  Matches  121;  Conservative    0;  Mismatches  132;  Indels    0;  Gaps    0;

Qy        401 ATCAGTCTGCTTTACAACAGCCCGCCGCGTTGGCGTTTTCGGGATCATCCCCGCAACACC 460
              ||||| || |   ||    || ||    |  || | |  | |  |       |||| |||
Db        769 ATCAGCCTTCCCAACGCAGGCACGGGCAGCCGGGGCTGACAGACTGTAGAGCGCAGCACG 828

Qy        461 CCCCACCTCAAACGACGTCGGCATCCGTTGGACAACAGCAACACGTGGTGTCGGGGTCTT 520
              | |||    | |  ||| || | ||     |  | || ||||       |  |||| | |
Db        829 CTCCATGACCGAGCACGCCGCCGTCTCCGGCGCGACCGCAAGGGTCTGCCTCGGCGGCGC 888

Qy        521 CTGGACAACAACCGCAACAGGGAGCACAGTCAAGCACTGTCCAGCCAACAACCGGATCAC 580
              | | || |  | ||  |  |  || | | | ||| |    | || | |     |  |
Db        889 GTCGCCATCGTCAGCCCCGCGTCGCGCGGCCGAGCCCAACGCCGCTAGCCCGGCGCACCT 948

Qy        581 CGCCCGCGGCCCAAGGCGTGCCACAGTCTACCCCGCCCCCAACCCAAAATACCCCCCAGG 640
              |   |||| | |      | |   ||  | ||||   |||||||||   |  | || | |
Db        949 CCGGCGCGCCTCCGCCAGCGTACCACGCCACCCGCACCCCAACCCGTGAGCCGCCACTGC 1008

Qy        641 GGGGTAAGGGACA 653
                | | ||  ||
Db       1009 CCTGCACGGTGCA 1021
```

```
RESULT 7
AG135197
LOCUS       AG135197               710 bp    DNA     linear   GSS 04-NOV-2001
DEFINITION  Pan troglodytes DNA, clone: PTB-148G11.F, genomic survey sequence.
ACCESSION   AG135197
VERSION     AG135197.1  GI:16664875
KEYWORDS    GSS.
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE   1
  AUTHORS   Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
            Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     BAC end sequences of Library PTB
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 710)
  AUTHORS   Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
            Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            (E-mail:chimpbes@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
            Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT     Clones are derived from the chimpanzee BAC library PTB This BAC end
            was generated during the R&D process and may have higher chance of
            clone tracking errors.
             PRIMERS
                 Sequencing: -21M13
             LIBRARY
                 Vector    : pKS145
                 R.Site 1  : SacI
                 R.Site 2  : SacI.
FEATURES             Location/Qualifiers
     source          1. .710
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9598"
                     /clone="PTB-148G11.F"
                     /sex="male"
                     /cell_type="lymphoblast"
                     /clone_lib="PTB Chimpanzee Male BAC Library"
ORIGIN

  Query Match             5.8%;  Score 40.8;  DB 9;  Length 710;
  Best Local Similarity   48.3%;  Pred. No. 0.92;
  Matches  114;  Conservative    0;  Mismatches  122;  Indels    0;  Gaps    0;

Qy        356 GAATCTCGGGGCCTACTATATTGCCGCAAAAAACACAGAGCGCCGATCAGTCTGCTTTAC 415
              ||| | |  | || || | |   || | || | ||   | |  |   ||    ||   ||
Db        394 GAACCCCCCGACCCACAAGAGGCCCCCCAACACCAGCAAACCACCCCCACAGAGCAGCAC 453

Qy        416 AACAGCCCGCCGCGTTGGCGTTTTCGGGATCATCCCCGCAACACCCCCCACCTCAAACGA 475
              |  | ||| || |      ||    || ||      | | | |||| |||| || ||  | |
Db        454 ACGACCCCCCCACCCCACCGCACACGCGACACAACACTCCACACACCCCCCCCCACCCCA 513

Qy        476 CGTCGGCATCCGTTGGACAACAGCAACACGTGGTGTCGGGGTCTTCTGGACAACAACCGC 535
              |||||  ||    ||  |    | |  |     |    |  | || ||||||| |
Db        514 ACACGGCAAACGCCCCACGCCCCGCCCCCCCGCCCCCACCCACGCCCCGAGAACAACCAC 573

Qy        536 AACAGGGAGCACAGTCAAGCACTGTCCAGCCAACAACCGGATCACCGCCCGCGGCC 591
              | ||  | |||  |     | |     ||| | | ||  || |   || |||  | ||
Db        574 AGCAACGCGCAAACGACAACCAAACCCAACGACCACACGAAAGCCCACCCCAGACC 629


RESULT 8
AG075888
LOCUS       AG075888               1156 bp    DNA     linear   GSS 03-NOV-2001
DEFINITION  Pan troglodytes DNA, clone: PTB-068O20.R, genomic survey sequence.
ACCESSION   AG075888
```

```
VERSION     AG075888.1  GI:16627690
KEYWORDS    GSS.
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE   1
  AUTHORS   Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
            Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     BAC end sequences of Library PTB
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1156)
  AUTHORS   Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
            Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            (E-mail:chimpbes@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
            Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT     Clones are derived from the chimpanzee BAC library PTB This BAC end
            was generated during the R&D process and may have higher chance of
            clone tracking errors.
             PRIMERS
                Sequencing: M13Rev
             LIBRARY
                Vector    : pKS145
                R.Site 1  : SacI
                R.Site 2  : SacI.
FEATURES             Location/Qualifiers
     source          1. .1156
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9598"
                     /clone="PTB-068020.R"
                     /sex="male"
                     /cell_type="lymphoblast"
                     /clone_lib="PTB Chimpanzee Male BAC Library"
ORIGIN

  Query Match             5.7%;  Score 40.2;  DB 9;  Length 1156;
  Best Local Similarity   50.8%;  Pred. No. 1.5;
  Matches   96;  Conservative    0;  Mismatches   93;  Indels    0;  Gaps    0;

Qy        449 CCCCGCAACACCCCCCACCTCAAACGACGTCGGCATCCGTTGGACAACAGCAACACGTGG 508
              | ||||  ||| ||| | | ||  |  | | || |||     |       ||  | ||  |
Db        328 CGCCGCCCCACGCCCGAACACAGCCCCCCGCAGCCTCCCGCCGGGCGGCGCCGCCCGCCG 387

Qy        509 TGTCGGGGTCTTCTGGACAACAACCGCAACAGGGAGCACAGTCAAGCACTGTCCAGCCAA 568
              | |||   |      |    ||| ||  |||       | |  || | |    |||||
Db        388 AGCCGGCAGCGCACCCAGGCAAACAGCCCCAGACCCGCCCGCAAATCCGCGCGGCGCCAA 447

Qy        569 CAACCGGATCACCGCCCGCGGCCCAAGGCGTGCCACAGTCTACCCCGCCCCCAACCCAAA 628
              |  || |  | |||||||||   |  |  |  || || | |  | || ||||||||||
Db        448 CCCCCCGCGCCCCGCCCGCCCACGCACACACGCAACCGGCCCCACCCCCCCCAACCCGCC 507

Qy        629 ATACCCCCC 637
              |  |||||
Db        508 GTGACCCCC 516


RESULT 9
CNS0169V
LOCUS       CNS0169V                1201 bp    DNA     linear   GSS 26-JUL-1999
DEFINITION  Drosophila melanogaster genome survey sequence SP6 end of BAC
            BACN15C06 of DrosBAC library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL106477
VERSION     AL106477.1  GI:5622081
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster (fruit fly)
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 1201)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the European Drosophila Genome Project (EDGP) -
            http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
            library (Dros BAC) was made by Alain Billaud at CBPH (Centre
            d'Etude du Polymorphisme Humain) with funding provided by a MRC
            project grant. The DNA was prepared from embryos by Alain Bucheton
            and Genevieve Payan. It has been constructed in the vector
            pBeloBAC11.
FEATURES             Location/Qualifiers
     source          1. .1201
                     /organism="Drosophila melanogaster"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7227"
                     /clone="BACN15C06"
                     /clone_lib="DrosBAC"
                     /plasmid="pBeloBAC11"
                     /note="end : SP6"
ORIGIN

  Query Match             5.7%;  Score 40.2;  DB 9;  Length 1201;
  Best Local Similarity   31.3%;  Pred. No. 1.6;
  Matches   92;  Conservative   59;  Mismatches  143;  Indels    0;  Gaps    0;

Qy        382 CAAAAAACACAGAGCGCCGATCAGTCTGCTTTACAACAGCCCGCCGCGTTGGCGTTTTCG 441
              |||||  :||  | | || |          :   | |:  || :| :::|  :| |      |
Db        841 CAAAACMCAGGGCGRGCGGGAGGTAGCASSGGAAAWGGGCGVGAMSSGGCSGGGGGMGGG 900

Qy        442 GGATCATCCCCGCAACACCCCCCACCTCAAACGACGTCGGCATCCGTTGGACAACAGCAA 501
              :|:|     :::|  :  ::::: |: :   |      :| |·| :|   |: || |   |
Db        901 SGRTGCMAAMMSCCSSSGSSSSSGCSGSSGCCCSGHGMGWCGTNMGAAAGWGAAAAAASA 960

Qy        502 CACGTGGTGTCGGGGTCTTCTGGACAACAACCGCAACAGGGAGCACAGTCAAGCACTGTC 561
              || | :   |  :|   ||  || ||  | : |::| | | :   || :   : |
Db        961 AACMTARGAGGGAAHYCGAGAGGGGAAAAAGKGGTMGARRGGGGARRAGAAAMMSAMSAC 1020

Qy        562 CAGCCAACAACCGGATCACCGCCCGCGGCCCAAGGCGTGCCACAGTCTACCCCGCCCCCA 621
              ||  :||  ::: :  :|||::     :||  :  :  |:      |   ||| :||||:
Db       1021 CAMAMAAAYMSVMVCCMACCSSGAAAMSCCGCMMCSCAGMACACCSCACACCCCMCCCCM 1080

Qy        622 ACCCAAAATACCCCCCAGGGGGGTAAGGGACAGACCTTGTCACACACGGGACAA 675
              ||:|::    ||||:|  : |   |: ||:|: :  :  |  :||||:   |::|
Db       1081 ACMCRMCCCCCCCCVCGCSCGCAGAMAGGMCSSSRSGVGGAWCACASMCCAMVA 1134


RESULT 10
CNS006ON/c
LOCUS       CNS006ON                910 bp    DNA     linear   GSS 03-JUN-1999
DEFINITION  Drosophila melanogaster genome survey sequence T7 end of BAC #
            BACR14J21 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL065629
VERSION     AL065629.1  GI:4944698
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster (fruit fly)
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 910)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
```

collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

```
FEATURES             Location/Qualifiers
     source          1. .910
                     /organism="Drosophila melanogaster"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7227"
                     /clone="BACR14J21"
                     /clone_lib="RPCI-98"
                     /note="end : T7"
ORIGIN
```

```
  Query Match             5.7%;  Score 40;  DB 9;  Length 910;
  Best Local Similarity   17.4%;  Pred. No. 1.7;
  Matches   63;  Conservative  141;  Mismatches  158;  Indels    0;  Gaps    0;

Qy        276 AAGACCAACGGTTGGCTTAAAACGTACGTTTAACCCGCGTATTATACGACCACAACCCCC 335
              ||:|:: |::   ::: : ::: |:    | :: : : :    : :|:    |:::::
Db        852 AASASSSAVSACGSSSSSVSSVVSSASASASAGVSRSASSVSVSSVGSASRSRSASVSSS 793

Qy        336 AAATCCATCCATGAGTTTGGGAATCTCGGGGCCTACTATATTGCCGCAAAAAACACAGAG 395
              | : | : :  |    : :|: : :: ::: :  |:   :: ::||::|::|:| ::
Db        792 SSAGSAAVSAVAMACASMVASAVASVMSASVVAARASAKSAASMARVAAVVARVASAAVV 733

Qy        396 CGCCGATCAGTCTGCTTTACAACAGCCCGCCGCGTTGGCGTTTTCGGGATCATCCCCGCA 455
              : : |  :  : ::   |::::|:  :: :::: : ::    : ::| :: :| : :
Db        732 ASASAAVAVAAVSRSVASASRVMASAGSSASSSSSSSASSSASSSASSASMVVSCRSAVS 673

Qy        456 ACACCCCCCACCTCAAACGACGTCGGCATCCGTTGGACAACAGCAACACGTGGTGTCGGG 515
              :    :          :|: ::   :    : :  :| : :: :: :| |: :: : : ::
Db        672 VASAASAAASARSASRASAVMAVAVAASSSAVASSGSSSVSCSSSRCGCSASSSSASASS 613

Qy        516 GTCTTCTGGACAACAACCGCAACAGGGAGCACAGTCAAGCACTGTCCAGCCAACAACCGG 575
              | :  : :: :  |  | :|| : ::: :| | : :  :: : | :: ::: :   |::
Db        612 GSSSSSSSSSSCGCGSCASCASSSSRSGSCGCGSSVGCVSSSCGSSSSSSSSSSSSGCSS 553

Qy        576 ATCACCGCCCGCGGCCCAAGGCGTGCCACAGTCTACCCCGCCCCCAACCCAAAATACCCC 635
              | ||:|:| | :|||   ::    | |:: :  ::|| ||||  :|| :|:|:||||:
Db        552 SSCCCCSCSCACCRCCCCCCRSCSCGCCCMSSSCCSSCCCCCCCCASMCCAWAMAHACCCS 493

Qy        636 CC 637
              ||
Db        492 CC 491
```

```
RESULT 11
CE497548
LOCUS       CE497548                454 bp    DNA     linear   GSS 28-SEP-2003
DEFINITION  tigr-gss-dog-17000327311670 Dog Library Canis familiaris genomic,
            genomic survey sequence.
ACCESSION   CE497548
VERSION     CE497548.1  GI:36814329
KEYWORDS    GSS.
SOURCE      Canis familiaris (dog)
  ORGANISM  Canis familiaris
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE   1  (bases 1 to 454)
  AUTHORS   Kirkness,E.F., Bafna,V., Halpern,A.L., Levy,S., Remington,K.,
            Rusch,D.B., Delcher,A.L., Pop,M., Wang,W., Fraser,C.M. and
            Venter,J.C.
  TITLE     The dog genome: survey sequencing and comparative analysis
  JOURNAL   Science 301 (5641), 1898-1903 (2003)
  MEDLINE   22875432
   PUBMED   14512627
COMMENT     Contact: Kirkness EF
            The Institute for Genomic Research
            Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
            Rockville, MD 20850, USA
            Tel: 301-838-0200
            Fax: 301-838-0208
            Email: ekirknes@tigr.org
            Class: shotgun.
FEATURES             Location/Qualifiers
     source          1. .454
                     /organism="Canis familiaris"
                     /mol_type="genomic DNA"
                     /strain="Standard Poodle"
                     /db_xref="taxon:9615"
                     /clone_lib="Dog Library"
                     /note="Site_1: BstXI; Libraries were prepared from
                     peripheral blood"
ORIGIN
```

```
  Query Match             5.6%;  Score 39.6;  DB 9;  Length 454;
  Best Local Similarity   49.5%;  Pred. No. 1.9;
  Matches  102;  Conservative    0;  Mismatches  104;  Indels    0;  Gaps    0;

Qy        479 CGGCATCCGTTGGACAACAGCAACACGTGGTGTCGGGGTCTTCTGGACAACAACCGCAAC 538
              |||   ||| ||||| |||||  |  | ||  ||  || |  | || |  |  || |
Db        237 CGGTGGCCGGTGGACCACAGCTTCTAGAGGGCTCCAGGACCCCAGGGCTCCTCCCACTGG 296

Qy        539 AGGGAGCACAGTCAAGCACTGTCCAGCCAACAACCGGATCACCGCCCGCGGCCCAAGGCG 598
              ||||  |      ||  ||| ||| ||  | | || ||   | | |||    | |
Db        297 CCCCCTGACAGCCTTCTCCTCCCCACCCACCAGGCTGGTCCCCTAGCTCAGCCTGCAGGG 356

Qy        599 TGCCACAGTCTACCCCGCCCCCAACCCAAAATACCCCCCAGGGGGGTAAGGGACAGACCT 658
              | ||  | |  || | | || ||  || |   |  ||| |       || ||   |||
Db        357 GGTCAGGGCCGCCCACTCACTCACCCTGAAGTGAGCGGCAGTGACCCGGGGCACCTGCCT 416

Qy        659 TGTCACACACGGGACAATCTGGAAAC 684
                ||||  |  |      ||||| |
Db        417 GTGCACAGCCCTGGGTGCATGGAACC 442
```

```
RESULT 12
BZ144309
LOCUS       BZ144309                1055 bp    DNA     linear   GSS 11-OCT-2002
DEFINITION  CH230-462F7.TJ CHORI-230 Segment 2 Rattus norvegicus genomic clone
            CH230-462F7, genomic survey sequence.
ACCESSION   BZ144309
VERSION     BZ144309.1  GI:23785256
KEYWORDS    GSS.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 1055)
  AUTHORS   Zhao,S., Shetty,J., Shatsman,S., Tsegaye,G., Geer,K.,
            Shvartsbeyn,A., Gebregeorgis,E., Overton,L., Russell,D., Chen,D.,
            Riggs,F., de Jong,P. and Fraser,C.M.
  TITLE     Rat BAC End Sequences from Library CHORI-230 MboI segment
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: szhao@tigr.org
            Clones are derived from the rat BAC library CHORI-230
            (http://www.chori.org/bacpac/rat230.htm). For BAC library
            availability, please contact Pieter de Jong (pdejong@mail.cho.org).
```

```
            Clones may be purchased from BACPAC Resources
            (http://www.chori.org/bacpac/or ering_information.htm). BAC end
            page: http://www.tigr.org/tdb/bac_ends/rat/bac_end_intro.html
            Plate: 462  row: F  column: 7
            Seq primer: SP6
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .1055
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /strain="BN/SsNHsd/MCW"
                     /db_xref="taxon:10116"
                     /clone="CH230-462F7"
                     /sex="Female"
                     /cell_type="Brain"
                     /clone_lib="CHORI-230 Segment 2"
                     /note="Vector: pTARBAC1.3; Site_1: MboI; Site_2: MboI;
                     CHORI-230 Rat (BN/SsNHsd/MCW) BAC library produced by
                     Pieter de Jong"
ORIGIN

  Query Match             5.6%;  Score 39.6;  DB 8;  Length 1055;
  Best Local Similarity   48.6%;  Pred. No. 2.3;
  Matches  108;  Conservative    0;  Mismatches  114;  Indels    0;  Gaps    0;

Qy        406 TCTGCTTTACAACAGCCCGCCGCGTTGGCGTTTTCGGGATCATCCCCCGCAACACCCCCCA 465
              || | |   | ||| || |||  | |    || ||   |||||| || | ||  || |||
Db        514 TCAGTTGAGCTACATCCAGCCCTGCTTCATTTATCATCATCATCACCACCACCACCACCA 573

Qy        466 CCTCAAACGACGTCGGCATCCGTTGGACAACAGCAACACGTGGTGTCGGGGTCTTCTGGA 525
              || | | | | ||  || |       || ||  | ||          ||    || ||   |
Db        574 CCACCATCACCATCACCACCACCACCACTACCACCACCACCATCATCATCATCATCA 633

Qy        526 CAACAACCGCAACAGGGAGCACAGTCAAGCACTGTCCAGCCAACAACCGGATCACCGCCC 585
                |  ||| | ||    | ||   |||    |      ||| || |   ||||||
Db        634 TCATCACCACCACCACCATCATCATCACCACCACCATCACCATCATCATCATCACCACCA 693

Qy        586 GCGGCCCAAGGCGTGCCACAGTCTACCCCGCCCCCAACCCAA 627
                |  | | |      ||||  ||  |  | || ||| | | |
Db        694 CCACCACCACCACCACCACCATCATCATCACCACCACCACCA 735


RESULT 13
AG471975/c
LOCUS       AG471975                870 bp    DNA     linear   GSS 04-JUN-2004
DEFINITION  Mus musculus molossinus DNA, clone:MSMg01-363J02.TJ, genomic survey
            sequence.
ACCESSION   AG471975
VERSION     AG471975.1  GI:48179205
KEYWORDS    GSS.
SOURCE      Mus musculus molossinus
  ORGANISM  Mus musculus molossinus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
  TITLE     BAC end Sequences of Library MSMg01
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 870)
  AUTHORS   Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            (E-mail:hattori@gsc.riken.jp, URL:http://hgp.gsc.riken.go.jp/,
            Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT     Clones are derived from the mouse BAC library MSMg01. For BAC
            library availability, please contact Kuniya Abe (abe@rtc.riken.jp).
            Tsukuba Institude, Bio Resource Center,
            The Institute of Physical and Chemical Research (RIKEN) 3-1-1
            Koyadai, Tsukuba, 305-0074 Japan
            phone: 81-298-36-9189, fax: 81-298-36-9199
            e-mail: abe@rtc.riken.jp
            PRIMERS
            Sequencing : TJ
            LIBRARY
            Vector     : pBACe3.6
            R.Site 1   : EcoRI
            R.Site 2   : EcoRI.
FEATURES             Location/Qualifiers
     source          1. .870
                     /organism="Mus musculus molossinus"
                     /mol_type="genomic DNA"
                     /sub_species="molossinus"
                     /db_xref="taxon:57486"
                     /clone="MSMg01-363J02.TJ"
                     /sex="male"
                     /tissue_type="mixture of kidney and spleen"
                     /clone_lib="MSMg01 Mouse Male BAC Library"
ORIGIN

  Query Match             5.6%;  Score 39.4;  DB 9;  Length 870;
  Best Local Similarity   44.6%;  Pred. No. 2.5;
  Matches  115;  Conservative    0;  Mismatches  143;  Indels    0;  Gaps    0;

Qy        379 CCGCAAAAAACACAGAGCGCCGATCAGTCTGCTTTACAACAGCCCGCCGCGTTGGCGTTT 438
              ||||| || ||||| |       |  | | ||    |  |   | |  | |  || |
Db        506 CCGCANAACACACACACGCGAAAGAGGGCGGCACCCCGCCCCGCGGGGGGGGGGGGGGGG 447

Qy        439 TCGGGATCATCCCCGCAACACCCCCCACCTCAAACGACGTCGGCATCCGTTGGACAACAG 498
              ||||  || |||  |  |  ||||| || |   |  | |  | |    || || | |
Db        446 GGGGGAACAACCCGCCCCCNNCCCCCCCCCCCCCGCCNCCCCCCCCCNCAAAAGGNCACCGG 387

Qy        499 CAACACGTGGTGTCGGGGTCTTCTGGACAACAACCGCAACAGGGAGCACAGTCAAGCACT 558
                | ||  |  ||||      || |  |  |  |||||     ||||  ||    |
Db        386 AGGGCGGGGGGGGGGGGGCGGGGCGGNCCCCCGCNCCAACACCACCCACACGCACCACCC 327

Qy        559 GTCCAGCCAACAACCGGATCACCGCCCGCGGCCCAAGGCGTGCCACAGTCTACCCCGCCC 618
                |       |||  |     || |||||    |    |   | | | | |||
Db        326 ACNNCCNCNNNNNGGGGAAAAAAAGGCGGGGCCCCGCCCCACACCACGACNCCACGCCCC 267

Qy        619 CCAACCCAAAATACCCCC 636
              ||  | ||| | |||| |
Db        266 CCCCCACAACACACCCAC 249


RESULT 14
CNS0073W/c
LOCUS       CNS0073W                922 bp    DNA     linear   GSS 03-JUN-1999
DEFINITION  Drosophila melanogaster genome survey sequence TET3 end of BAC #
            BACR14D09 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL066784
VERSION     AL066784.1  GI:4945247
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster (fruit fly)
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 922)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see  http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutoyo Osoegawa and
            Aaron Mammoser in Pieter de Jong's laboratory in the Department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RPCI-98 and was constructed by partial
            EcoRI digestion of Drosophila DNA provided by the BDGP from the
```

isogenic strain y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES Location/Qualifiers
source 1. .922
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR14D09"
/clone_lib="RPCI-98"
/note="end : TET3"

ORIGIN

Query Match 5.6%; Score 39.4; DB 9; Length 922;
Best Local Similarity 19.2%; Pred. No. 2.5;
Matches 71; Conservative 123; Mismatches 176; Indels 0; Gaps 0;

```
Qy        325 CCACAACCCCCAAATCCATCCATGAGTTTGGGAATCTCGGGGCCTACTATATTGCCGCAA 384
              :: | :|||||::: ::  ::           :: : :    |: :: : :   |   ::
Db        916 MMCCCMCCCCCMMMCMMCMMMCMCCMCMMMCMMMCMMMMMCMCMCMMMMMMMMACACAMM 857

Qy        385 AAAACACAGAGCGCCGATCAGTCTGCTTTACAACAGCCCGCCGCGTTGGCGTTTTCGGGA 444
              ||: :::| : : |: | ::  |  :   ::|: : ::: :: :     |          :
Db        856 AAMNMMMACMMMMCMMACMMAMCCMMACMMMAMAMMMMMMMMAMMAMCACMAMMMACACM 797

Qy        445 TCATCCCCGCAACACCCCCCACCTCAAACGACGTCGGCATCCGTTGGACAACAGCAACAC 504
                : ::|: :::::::|::|::|: || :: ||  :  |  :|     ||: ::  :| ::
Db        796 CAMMMMCMMMMMMMMCMMCMMCMCCACMMMACACMAMCCMMCMMCMMACMCMMAAMAAMM 737

Qy        505 GTGGTGTCGGGGTCTTCTGGACAACAACCGCAACAGGGAGCACAGTCAAGCACTGTCCAG 564
                        :      ::||::| : ::|::     :| :   :: |::   | :
Db        736 MMACAMMAMAAMMMMAMAAMMAAMMAAMMMMAMMCCMCCMAAMAMAMMACMMCMMCAMM 677

Qy        565 CCAACAACCGGATCACCGCCCGCGGCCCAAGGCGTGCCACAGTCTACCCCGCCCCCAACC 624
              ::::||:::  | :|:: |||    ::|:   :   : |:|  : ||::: ::  ::||:
Db        676 MMMMCAMMMMMAMMAMMMCCCAACAMMCMCACMMMCMAAMAMMMMACMMMAMMAAMMACM 617

Qy        625 CAAAATACCCCCCAGGGGGGTAAGGGACAGACCTTGTCACACACGGGACAATCTGGAAAC 684
              |||:: ||   ::|        |:  ||| ::     ::| ||:   |::| :   | ||
Db        616 CAAMMCACAAAMMACMCMMCMCASACACAMVMRAAMMMMCCCASAMAAMMAMMAAMACAC 557

Qy        685 GCTTCAAGAA 694
                |:||
Db        556 MAMCACASAA 547
```

RESULT 15
CC910437
LOCUS CC910437 787 bp DNA linear GSS 08-AUG-2003
DEFINITION t061h01ba.r1 TAMBT Bos taurus genomic clone t061h01ba, genomic survey sequence.
ACCESSION CC910437
VERSION CC910437.1 GI:33532793
KEYWORDS GSS.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 787)
AUTHORS Lin,S., Najar,F.Z., Adelson,D., Gill,C.A. and Roe,B.A.
TITLE Bovine BAC End Sequences from Library TAMBT
JOURNAL Unpublished (2003)
COMMENT Contact: Bruce A. Roe
Advanced Center for Genome Technology
University of Oklahoma Department of Chemistry and Biochemistry
620 Parrington Oval, Room 208, Norman, OK 73019, USA
Tel: 405 325 4912
Fax: 405 325 7762
Email: broe@ou.edu
Class: BAC ends

High quality sequence start: 8
High quality sequence stop: 311.

FEATURES Location/Qualifiers
source 1. .787
/organism="Bos taurus"
/mol_type="genomic DNA"
/strain="Angus bull T A M U Shoshone Y6 11519666"
/db_xref="taxon:9913"
/clone="t061h01ba"
/sex="Male"
/cell_type="Blood"
/clone_lib="TAMBT"
/note="Vector: pBeloBAC11; Site_1: HindIII; Site_2: HindIII; TAMBT Bovine BAC library (Male) produced by Texas A&M University, Department of Animal Science."

ORIGIN

Query Match 5.6%; Score 39.2; DB 9; Length 787;
Best Local Similarity 50.5%; Pred. No. 2.8;
Matches 95; Conservative 0; Mismatches 93; Indels 0; Gaps 0;

```
Qy        449 CCCCGCAACACCCCCCACCTCAAACGACGTCGGCATCCGTTGGACAACAGCAACACGTGG 508
              |||| |  | ||| ||    |   | ||  |  |  ||      |  |  |  | |  |
Db        574 CCCCACCCCGCCCACCCAGCCCCCCCACCGCCACCGCCACACCCCCCCCCCCCCCCGGC 633

Qy        509 TGTCGGGGTCTTCTGGACAACAACCGCAACAGGGAGCACAGTCAAGCACTGTCCAGCCAA 568
              | |   | | |   ||| |  || |||||    ||  || ||| | | | ||  |||
Db        634 CGCCACCGCCCCCCCCACACCCCCCCCAACAACCCGCCAAGACAACCCCCGGCCCCCCAC 693

Qy        569 CAACCGGATCACCGCCCGCGGCCCAAGGCGTGCCACAGTCTACCCCGCCCCCAACCCAAA 628
              |  ||  | | || ||| |  |||    |  || |   |  || || |||| ||||| |
Db        694 CCCCCCCAACGCCCCCCCCCCCCCCCCCCCCCCCCCCGAGCCCCCACGGCCCCCACCCACA 753

Qy        629 ATACCCCC 636
              | ||| ||
Db        754 ACACCGCC 761
```

Search completed: September 25, 2005, 23:45:38
Job time : 3224 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: September 25, 2005, 21:34:07 ; Search time 181 Seconds
(without alignments)
6373.343 Million cell updates/sec

Title: US-09-874-140-1
Perfect score: 705
Sequence: 1 atgacacaacccgcatcgtc.....cttcaagaagtcgtagggtg 705

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	705	100.0	705	3	US-09-219-337-1	Sequence 1, Appli
c	2	705	100.0	124884	4	US-09-661-596A-76	Sequence 76, Appl
c	3	705	100.0	124884	4	US-09-913-514-1	Sequence 1, Appli
c	4	703.4	99.8	125157	4	US-09-913-514-2	Sequence 2, Appli
	5	36.2	5.1	4403765	3	US-09-103-840A-2	Sequence 2, Appli
	6	36.2	5.1	4411529	3	US-09-103-840A-1	Sequence 1, Appli
c	7	35.8	5.1	4403765	3	US-09-103-840A-2	Sequence 2, Appli
c	8	35.8	5.1	4411529	3	US-09-103-840A-1	Sequence 1, Appli
c	9	35	5.0	601	4	US-09-949-016-160829	Sequence 160829,
	10	35	5.0	45427	4	US-09-949-016-16243	Sequence 16243, A
	11	34	4.8	642	3	US-08-911-853-12	Sequence 12, Appl
	12	34	4.8	642	3	US-09-479-409-12	Sequence 12, Appl
	13	34	4.8	642	3	US-09-479-453-12	Sequence 12, Appl
c	14	34	4.8	17612	3	US-08-911-853-29	Sequence 29, Appl
c	15	34	4.8	17612	3	US-09-479-409-29	Sequence 29, Appl
c	16	34	4.8	17612	3	US-09-479-453-29	Sequence 29, Appl
	17	33.6	4.8	505	4	US-09-621-976-15639	Sequence 15639, A
	18	33.6	4.8	34230	4	US-09-949-016-12052	Sequence 12052, A
	19	33.6	4.8	128470	4	US-09-949-016-13765	Sequence 13765, A
c	20	33.4	4.7	8399	4	US-09-949-016-12353	Sequence 12353, A
c	21	33.4	4.7	8625	3	US-08-980-832-1	Sequence 1, Appli
c	22	33.4	4.7	8625	4	US-09-920-923B-1	Sequence 1, Appli
c	23	33.4	4.7	11233	3	US-08-980-832-27	Sequence 27, Appl
c	24	33.4	4.7	11233	4	US-09-920-923B-27	Sequence 27, Appl
	25	33.2	4.7	486	4	US-09-252-991A-7451	Sequence 7451, Ap
	26	33	4.7	1179	4	US-09-489-039A-7127	Sequence 7127, Ap
c	27	32.6	4.6	43280	2	US-08-804-227C-1	Sequence 1, Appli
c	28	32.2	4.6	711	4	US-09-902-540-3794	Sequence 3794, Ap
c	29	32.2	4.6	18551	4	US-09-902-540-1187	Sequence 1187, Ap
c	30	32	4.5	28438	4	US-09-820-790B-3	Sequence 3, Appli
	31	31.8	4.5	2542	4	US-09-902-540-6580	Sequence 6580, Ap
	32	31.8	4.5	3312	4	US-09-902-540-493	Sequence 493, App
	33	31.8	4.5	109690	4	US-09-949-016-13525	Sequence 13525, A
	34	31.6	4.5	1767	4	US-09-252-991A-2290	Sequence 2290, Ap
	35	31.6	4.5	1887	4	US-09-252-991A-2209	Sequence 2209, Ap
c	36	31.6	4.5	2250	4	US-09-252-991A-2623	Sequence 2623, Ap
c	37	31.6	4.5	4200	1	US-07-841-654B-1	Sequence 1, Appli
c	38	31.6	4.5	4200	1	US-07-946-234A-1	Sequence 1, Appli
c	39	31.6	4.5	4200	1	US-08-123-161A-1	Sequence 1, Appli
c	40	31.6	4.5	4200	1	US-08-483-278-1	Sequence 1, Appli
c	41	31.6	4.5	4200	5	PCT-US93-01560-1	Sequence 1, Appli
	42	31.6	4.5	18558	4	US-09-949-016-13688	Sequence 13688, A
c	43	31.6	4.5	152132	4	US-09-949-016-13845	Sequence 13845, A
c	44	31.6	4.5	152145	4	US-09-949-016-12371	Sequence 12371, A
c	45	31.4	4.5	485	4	US-09-949-016-198725	Sequence 198725,

ALIGNMENTS

RESULT 1
US-09-219-337-1
; Sequence 1, Application US/09219337
; Patent No. 6258363
; GENERAL INFORMATION:
; APPLICANT: EICKMANN, Markus
; APPLICANT: GICKLHORN, Dorothee
; APPLICANT: RADSAK, Klaus
; APPLICANT: HAUSER, Hans-Peter
; APPLICANT: GIESENDORF, Bernhard
; TITLE OF INVENTION: VARICELLA ZOSTER VIRUS (VZV) IMMUNOREACTIVE PROTEIN
; TITLE OF INVENTION: VP26 AND ITS DIAGNOSTIC USE
; FILE REFERENCE: 058315/0127
; CURRENT APPLICATION NUMBER: US/09/219,337
; CURRENT FILING DATE: 1998-12-23
; EARLIER APPLICATION NUMBER: DE 197 57 765.2
; EARLIER FILING DATE: 1997-12-23
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 705
; TYPE: DNA
; ORGANISM: Varicella Zoster Virus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(705)
US-09-219-337-1

Query Match 100.0%; Score 705; DB 3; Length 705;
Best Local Similarity 100.0%; Pred. No. 1.1e-227;
Matches 705; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ATGACACAACCCGCATCGTCTCGTGTAGTCTTTGATCCCAGCAACCCCACCACATTTTCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGACACAACCCGCATCGTCTCGTGTAGTCTTTGATCCCAGCAACCCCACCACATTTTCG 60

Qy         61 GTGGAAGCAATTGCGGCTTACACCCCCGTTGCTTTAATACGACTTTTAAACGCCAGTGGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTGGAAGCAATTGCGGCTTACACCCCCGTTGCTTTAATACGACTTTTAAACGCCAGTGGA 120

Qy        121 CCTTTGCAACCTGGTCACCGTGTGGACATCGCTGATGCCAGAAGCATTTACACCGTGGGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCTTTGCAACCTGGTCACCGTGTGGACATCGCTGATGCCAGAAGCATTTACACCGTGGGA 180

Qy        181 GCCGCGGCCAGTGCCGCGCGTGCACGCGCTAACCATAATGCAAATACGATACGCCGAACG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCCGCGGCCAGTGCCGCGCGTGCACGCGCTAACCATAATGCAAATACGATACGCCGAACG 240

Qy        241 GCCATGTTTGCCGAGACTGACCCTATGACATGGTTAAGACCAACGGTTGGCTTAAAACGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        241 GCCATGTTTGCCGAGACTGACCCTATGACATGGTTAAGACCAACGGTTGGCTTAAAACGT 300

Qy        301 ACGTTTAACCCGCGTATTATACGACCACAACCCCCAAATCCATCCATGAGTTTGGGAATC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ACGTTTAACCCGCGTATTATACGACCACAACCCCCAAATCCATCCATGAGTTTGGGAATC 360

Qy        361 TCGGGGCCTACTATATTGCCGCAAAAAACACAGAGCGCCGATCAGTCTGCTTTACAACAG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TCGGGGCCTACTATATTGCCGCAAAAAACACAGAGCGCCGATCAGTCTGCTTTACAACAG 420

Qy        421 CCCGCCGCGTTGGCGTTTTCGGGATCATCCCCGCAACACCCCCCACCTCAAACGACGTCG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CCCGCCGCGTTGGCGTTTTCGGGATCATCCCCGCAACACCCCCCACCTCAAACGACGTCG 480

Qy        481 GCATCCGTTGGACAACAGCAACACGTGGTGTCGGGGTCTTCTGGACAACAACCGCAACAG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCATCCGTTGGACAACAGCAACACGTGGTGTCGGGGTCTTCTGGACAACAACCGCAACAG 540

Qy        541 GGAGCACAGTCAAGCACTGTCCAGCCAACAACCGGATCACCGCCCGCGGCCCAAGGCGTG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GGAGCACAGTCAAGCACTGTCCAGCCAACAACCGGATCACCGCCCGCGGCCCAAGGCGTG 600

Qy        601 CCACAGTCTACCCCGCCCCCAACCCAAAATACCCCCCAGGGGGGTAAGGGACAGACCTTG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CCACAGTCTACCCCGCCCCCAACCCAAAATACCCCCCAGGGGGGTAAGGGACAGACCTTG 660

Qy        661 TCACACACGGGACAATCTGGAAACGCTTCAAGAAGTCGTAGGGTG 705
              |||||||||||||||||||||||||||||||||||||||||||||
Db        661 TCACACACGGGACAATCTGGAAACGCTTCAAGAAGTCGTAGGGTG 705
```

RESULT 2
US-09-661-596A-76/c
; Sequence 76, Application US/09661596A
; Patent No. 6528066
; GENERAL INFORMATION:
; APPLICANT: Grose, Charles
; APPLICANT: Santos, Richard
; TITLE OF INVENTION: VARIANT VARICELLA-ZOSTER VIRUSES AND METHODS OF USE
; FILE REFERENCE: 140.0011 0101
; CURRENT APPLICATION NUMBER: US/09/661,596A
; CURRENT FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: US 60/153,779
; PRIOR FILING DATE: 1999-09-14
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 76
; LENGTH: 124884
; TYPE: DNA
; ORGANISM: Varicella zoster
US-09-661-596A-76

Query Match 100.0%; Score 705; DB 4; Length 124884;
Best Local Similarity 100.0%; Pred. No. 1.7e-226;
Matches 705; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ATGACACAACCCGCATCGTCTCGTGTAGTCTTTGATCCCAGCAACCCCACCACATTTTCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      43138 ATGACACAACCCGCATCGTCTCGTGTAGTCTTTGATCCCAGCAACCCCACCACATTTTCG 43079

Qy         61 GTGGAAGCAATTGCGGCTTACACCCCCGTTGCTTTAATACGACTTTTAAACGCCAGTGGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      43078 GTGGAAGCAATTGCGGCTTACACCCCCGTTGCTTTAATACGACTTTTAAACGCCAGTGGA 43019

Qy        121 CCTTTGCAACCTGGTCACCGTGTGGACATCGCTGATGCCAGAAGCATTTACACCGTGGGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      43018 CCTTTGCAACCTGGTCACCGTGTGGACATCGCTGATGCCAGAAGCATTTACACCGTGGGA 42959

Qy        181 GCCGCGGCCAGTGCCGCGCGTGCACGCGCTAACCATAATGCAAATACGATACGCCGAACG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42958 GCCGCGGCCAGTGCCGCGCGTGCACGCGCTAACCATAATGCAAATACGATACGCCGAACG 42899

Qy        241 GCCATGTTTGCCGAGACTGACCCTATGACATGGTTAAGACCAACGGTTGGCTTAAAACGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42898 GCCATGTTTGCCGAGACTGACCCTATGACATGGTTAAGACCAACGGTTGGCTTAAAACGT 42839

Qy        301 ACGTTTAACCCGCGTATTATACGACCACAACCCCCAAATCCATCCATGAGTTTGGGAATC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42838 ACGTTTAACCCGCGTATTATACGACCACAACCCCCAAATCCATCCATGAGTTTGGGAATC 42779

Qy        361 TCGGGGCCTACTATATTGCCGCAAAAAACACAGAGCGCCGATCAGTCTGCTTTACAACAG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42778 TCGGGGCCTACTATATTGCCGCAAAAAACACAGAGCGCCGATCAGTCTGCTTTACAACAG 42719

Qy        421 CCCGCCGCGTTGGCGTTTTCGGGATCATCCCCGCAACACCCCCCACCTCAAACGACGTCG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42718 CCCGCCGCGTTGGCGTTTTCGGGATCATCCCCGCAACACCCCCCACCTCAAACGACGTCG 42659

Qy        481 GCATCCGTTGGACAACAGCAACACGTGGTGTCGGGGTCTTCTGGACAACAACCGCAACAG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42658 GCATCCGTTGGACAACAGCAACACGTGGTGTCGGGGTCTTCTGGACAACAACCGCAACAG 42599

Qy        541 GGAGCACAGTCAAGCACTGTCCAGCCAACAACCGGATCACCGCCCGCGGCCCAAGGCGTG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42598 GGAGCACAGTCAAGCACTGTCCAGCCAACAACCGGATCACCGCCCGCGGCCCAAGGCGTG 42539

Qy        601 CCACAGTCTACCCCGCCCCCAACCCAAAATACCCCCCAGGGGGGTAAGGGACAGACCTTG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42538 CCACAGTCTACCCCGCCCCCAACCCAAAATACCCCCCAGGGGGGTAAGGGACAGACCTTG 42479

Qy        661 TCACACACGGGACAATCTGGAAACGCTTCAAGAAGTCGTAGGGTG 705
              |||||||||||||||||||||||||||||||||||||||||||||
Db      42478 TCACACACGGGACAATCTGGAAACGCTTCAAGAAGTCGTAGGGTG 42434
```

RESULT 3
US-09-913-514-1/c
; Sequence 1, Application US/09913514
; Patent No. 6653069
; GENERAL INFORMATION:
; APPLICANT: GOMI, Yasuyuki
; APPLICANT: SUNAMACHI, Hiroki
; APPLICANT: TAKAHASHI, Michiaki
; APPLICANT: YAMANISHI, Koichi
; TITLE OF INVENTION: Method for Quality Control of an Attenuated Varicella Live Vaccine
; FILE REFERENCE: 0216-0454P
; CURRENT APPLICATION NUMBER: US/09/913,514
; CURRENT FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: PCT/JP01/00678
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: JP 2000-62734
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 124884
; TYPE: DNA
; ORGANISM: Varicella virus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(124884)
; OTHER INFORMATION: Dumas Strain
US-09-913-514-1

Query Match 100.0%; Score 705; DB 4; Length 124884;
Best Local Similarity 100.0%; Pred. No. 1.7e-226;
Matches 705; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ATGACACAACCCGCATCGTCTCGTGTAGTCTTTGATCCCAGCAACCCCACCACATTTTCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      43138 ATGACACAACCCGCATCGTCTCGTGTAGTCTTTGATCCCAGCAACCCCACCACATTTTCG 43079

Qy         61 GTGGAAGCAATTGCGGCTTACACCCCCGTTGCTTTAATACGACTTTTAAACGCCAGTGGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      43078 GTGGAAGCAATTGCGGCTTACACCCCCGTTGCTTTAATACGACTTTTAAACGCCAGTGGA 43019
```

```
Qy        121 CCTTTGCAACCTGGTCACCGTGTGGACATCGCTGATGCCAGAAGCATTTACACCGTGGGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      43018 CCTTTGCAACCTGGTCACCGTGTGGACATCGCTGATGCCAGAAGCATTTACACCGTGGGA 42959

Qy        181 GCCGCGGCCAGTGCCGCGCGTGCACGCGCTAACCATAATGCAAATACGATACGCCGAACG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42958 GCCGCGGCCAGTGCCGCGCGTGCACGCGCTAACCATAATGCAAATACGATACGCCGAACG 42899

Qy        241 GCCATGTTTGCCGAGACTGACCCTATGACATGGTTAAGACCAACGGTTGGCTTAAAACGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42898 GCCATGTTTGCCGAGACTGACCCTATGACATGGTTAAGACCAACGGTTGGCTTAAAACGT 42839

Qy        301 ACGTTTAACCCGCGTATTATACGACCACAACCCCCAAATCCATCCATGAGTTTGGGAATC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42838 ACGTTTAACCCGCGTATTATACGACCACAACCCCCAAATCCATCCATGAGTTTGGGAATC 42779

Qy        361 TCGGGGCCTACTATATTGCCGCAAAAAACACAGAGCGCCGATCAGTCTGCTTTACAACAG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42778 TCGGGGCCTACTATATTGCCGCAAAAAACACAGAGCGCCGATCAGTCTGCTTTACAACAG 42719

Qy        421 CCCGCCGCGTTGGCGTTTTCGGGATCATCCCCGCAACACCCCCCACCTCAAACGACGTCG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42718 CCCGCCGCGTTGGCGTTTTCGGGATCATCCCCGCAACACCCCCCACCTCAAACGACGTCG 42659

Qy        481 GCATCCGTTGGACAACAGCAACACGTGGTGTCGGGGTCTTCTGGACAACAACCGCAACAG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42658 GCATCCGTTGGACAACAGCAACACGTGGTGTCGGGGTCTTCTGGACAACAACCGCAACAG 42599

Qy        541 GGAGCACAGTCAAGCACTGTCCAGCCAACAACCGGATCACCGCCCGCGGCCCAAGGCGTG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42598 GGAGCACAGTCAAGCACTGTCCAGCCAACAACCGGATCACCGCCCGCGGCCCAAGGCGTG 42539

Qy        601 CCACAGTCTACCCCGCCCCCAACCCAAAATACCCCCCAGGGGGGTAAGGGACAGACCTTG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42538 CCACAGTCTACCCCGCCCCCAACCCAAAATACCCCCCAGGGGGGTAAGGGACAGACCTTG 42479

Qy        661 TCACACACGGGACAATCTGGAAACGCTTCAAGAAGTCGTAGGGTG 705
              |||||||||||||||||||||||||||||||||||||||||||||
Db      42478 TCACACACGGGACAATCTGGAAACGCTTCAAGAAGTCGTAGGGTG 42434
```

RESULT 4
US-09-913-514-2/c
; Sequence 2, Application US/09913514
; Patent No. 6653069
; GENERAL INFORMATION:
; APPLICANT: GOMI, Yasuyuki
; APPLICANT: SUNAMACHI, Hiroki
; APPLICANT: TAKAHASHI, Michiaki
; APPLICANT: YAMANISHI, Koichi
; TITLE OF INVENTION: Method for Quality Control of an Attenuated Varicella Live Vaccin
; FILE REFERENCE: 0216-0454P
; CURRENT APPLICATION NUMBER: US/09/913,514
; CURRENT FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: PCT/JP01/00678
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: JP 2000-62734
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 125157
; TYPE: DNA
; ORGANISM: Varicella virus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(125157)
; OTHER INFORMATION: Attenuated Oka strain
US-09-913-514-2

Query Match 99.8%; Score 703.4; DB 4; Length 125157;
Best Local Similarity 99.9%; Pred. No. 5.9e-226;
Matches 704; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 ATGACACAACCCGCATCGTCTCGTGTAGTCTTTGATCCCAGCAACCCCACCACATTTTCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      43113 ATGACACAACCCGCATCGTCTCGTGTAGTCTTTGATCCCAGCAACCCCACCACATTTTCG 43054

Qy         61 GTGGAAGCAATTGCGGCTTACACCCCCGTTGCTTTAATACGACTTTTAAACGCCAGTGGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      43053 GTGGAAGCAATTGCGGCTTACACCCCCGTTGCTTTAATACGACTTTTAAACGCCAGTGGA 42994

Qy        121 CCTTTGCAACCTGGTCACCGTGTGGACATCGCTGATGCCAGAAGCATTTACACCGTGGGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42993 CCTTTGCAACCTGGTCACCGTGTGGACATCGCTGATGCCAGAAGCATTTACACCGTGGGA 42934

Qy        181 GCCGCGGCCAGTGCCGCGCGTGCACGCGCTAACCATAATGCAAATACGATACGCCGAACG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42933 GCCGCGGCCAGTGCCGCGCGTGCACGCGCTAACCATAATGCAAATACGATACGCCGAACG 42874

Qy        241 GCCATGTTTGCCGAGACTGACCCTATGACATGGTTAAGACCAACGGTTGGCTTAAAACGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42873 GCCATGTTTGCCGAGACTGACCCTATGACATGGTTAAGACCAACGGTTGGCTTAAAACGT 42814

Qy        301 ACGTTTAACCCGCGTATTATACGACCACAACCCCCAAATCCATCCATGAGTTTGGGAATC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42813 ACGTTTAACCCGCGTATTATACGACCACAACCCCCAAATCCATCCATGAGTTTGGGAATC 42754

Qy        361 TCGGGGCCTACTATATTGCCGCAAAAAACACAGAGCGCCGATCAGTCTGCTTTACAACAG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42753 TCGGGGCCTACTATATTGCCGCAAAAAACACAGAGCGCCGATCAGTCTGCTTTACAACAG 42694

Qy        421 CCCGCCGCGTTGGCGTTTTCGGGATCATCCCCGCAACACCCCCCACCTCAAACGACGTCG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42693 CCCGCCGCGTTGGCGTTTTCGGGATCATCCCCGCAACACCCCCCACCTCAAACGACGTCG 42634

Qy        481 GCATCCGTTGGACAACAGCAACACGTGGTGTCGGGGTCTTCTGGACAACAACCGCAACAG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42633 GCATCCGTTGGACAACAGCAACACGTGGTGTCGGGGTCTTCTGGACAACAACCGCAACAG 42574

Qy        541 GGAGCACAGTCAAGCACTGTCCAGCCAACAACCGGATCACCGCCCGCGGCCCAAGGCGTG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42573 GGAGCACAGTCAAGCACTGTCCAGCCAACAACCGGATCACCGCCCGCGGCCCAAGGCGTG 42514

Qy        601 CCACAGTCTACCCCGCCCCCAACCCAAAATACCCCCCAGGGGGGTAAGGGACAGACCTTG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42513 CCACAGTCTACCCCGCCCCCAACCCAAAATACCCCCCAGGGGGGTAAGGGACAGACCTTG 42454

Qy        661 TCACACACGGGACAATCTGGAAACGCTTCAAGAAGTCGTAGGGTG 705
              || ||||||||||||||||||||||||||||||||||||||||||
Db      42453 TCCCACACGGGACAATCTGGAAACGCTTCAAGAAGTCGTAGGGTG 42409
```

RESULT 5
US-09-103-840A-2
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence

```
;   OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

  Query Match             5.1%;  Score 36.2;  DB 3;  Length 4403765;
  Best Local Similarity   45.6%;  Pred. No. 11;
  Matches  128;  Conservative    0;  Mismatches  153;  Indels    0;  Gaps    0;

Qy        397 GCCGATCAGTCTGCTTTACAACAGCCCGCCGCGTTGGCGTTTTCGGGATCATCCCCGCAA 456
              ||||| | | | ||    |    ||| || ||  | ||||| ||    || | ||
Db    2800579 GCCGACCCGCCAGCACCCCCGGCGCCGCCCACGCCGCCGTTTCCGCCGGCACCGCCATTG 2800638

Qy        457 CACCCCCCACCTCAAACGACGTCGGCATCCGTTGGACAACAGCAACACGTGGTGTCGGGG 516
              |  ||  | || |   |  || ||||  |||    ||        |   |  | |||
Db    2800639 CCGCCGGCGCCGCCGTCACCGCCGGCCGCCGAGATACCGGCCGGCCCATTTATTCCGGTA 2800698

Qy        517 TCTTCTGGACAACAACCGCAACAGGGAGCACAGTCAAGCACTGTCCAGCCAACAACCGGA 576
              | | | ||   |   | | | ||  |||| | ||    | | || ||| ||| |   |
Db    2800699 GCCCCGGCACCGCCGGCACCGCCGGCCGCACCGGCACCACCGGCCCCGCCGACACCGCCA 2800758

Qy        577 TCACCGCCCGCGGCCCAAGGCGTGCCACAGTCTACCCCGCCCCCAACCCAAAATACCCCC 636
              | || || ||| | ||       | |       | || |  ||| |        || ||
Db    2800759 ACGCCACCGGCGCCGCCGTTACCGAGAAGCCACCCTCCCCGACCACCGTTGCCGCCCACC 2800818

Qy        637 CAGGGGGGTAAGGGACAGACCTTGTCACACACGGGACAATC 677
              |    ||    |  |  | ||  ||| || | |   || |
Db    2800819 CCACCGGCACCGCCATCGCCCCCGTCCGACCCTGCCAAACC 2800859


RESULT 6
US-09-103-840A-1
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
;  APPLICANT: FLEISCHMAN, Robert D.
;  APPLICANT:  WHITE, Owen R.
;  APPLICANT:  FRASER, Claire M.
;  APPLICANT:  VENTER, John C.
;  TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
;  TITLE OF INVENTION:  TUBERCULOSIS
;  FILE REFERENCE: 24366-20007.00
;  CURRENT APPLICATION NUMBER: US/09/103,840A
;  CURRENT FILING DATE:  1998-06-24
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 4411529
;   TYPE: DNA
;   ORGANISM: Mycobacterium tuberculosis
;   OTHER INFORMATION: H37Rv
US-09-103-840A-1

  Query Match             5.1%;  Score 36.2;  DB 3;  Length 4411529;
  Best Local Similarity   45.6%;  Pred. No. 11;
  Matches  128;  Conservative    0;  Mismatches  153;  Indels    0;  Gaps    0;

Qy        397 GCCGATCAGTCTGCTTTACAACAGCCCGCCGCGTTGGCGTTTTCGGGATCATCCCCGCAA 456
              ||||| | | | ||    |    ||| || ||  | ||||| ||    || | ||
Db    2805126 GCCGACCCGCCAGCACCCCCGGCGCCGCCCACGCCGCCGTTTCCGCCGGCACCGCCATTG 2805185

Qy        457 CACCCCCCACCTCAAACGACGTCGGCATCCGTTGGACAACAGCAACACGTGGTGTCGGGG 516
              |  ||  | || |   |  || ||||  |||    ||        |   |  | |||
Db    2805186 CCGCCGGCGCCGCCGTCACCGCCGGCCGCCGAGATACCGGCCGGCCCATTTATTCCGGTA 2805245

Qy        517 TCTTCTGGACAACAACCGCAACAGGGAGCACAGTCAAGCACTGTCCAGCCAACAACCGGA 576
              | | | ||   |   | | | ||  |||| | ||    | | || ||| ||| |   |
Db    2805246 GCCCCGGCACCGCCGGCACCGCCGGCCGCACCGGCACCACCGGCCCCGCCGACACCGCCA 2805305

Qy        577 TCACCGCCCGCGGCCCAAGGCGTGCCACAGTCTACCCCGCCCCCAACCCAAAATACCCCC 636
              | || || ||| | ||       | |       | || |  ||| |        || ||
Db    2805306 ACGCCACCGGCGCCGCCGTTACCGAGAAGCCACCCTCCCCGACCACCGTTGCCGCCCACC 2805365

Qy        637 CAGGGGGGTAAGGGACAGACCTTGTCACACACGGGACAATC 677
              |    ||    |  |  | ||  ||| || | |   || |
Db    2805366 CCACCGGCACCGCCATCGCCCCCGTCCGACCCTGCCAAACC 2805406


RESULT 7
US-09-103-840A-2/c
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
;  APPLICANT: FLEISCHMAN, Robert D.
;  APPLICANT:  WHITE, Owen R.
;  APPLICANT:  FRASER, Claire M.
;  APPLICANT:  VENTER, John C.
;  TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
;  TITLE OF INVENTION:  TUBERCULOSIS
;  FILE REFERENCE: 24366-20007.00
;  CURRENT APPLICATION NUMBER: US/09/103,840A
;  CURRENT FILING DATE:  1998-06-24
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 4403765
;   TYPE: DNA
;   ORGANISM: Mycobacterium tuberculosis
;   FEATURE:
;   OTHER INFORMATION: CDC 1551
;   OTHER INFORMATION: "n" bases at various positions throughout the sequence
;   OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

  Query Match             5.1%;  Score 35.8;  DB 3;  Length 4403765;
  Best Local Similarity   52.3%;  Pred. No. 15;
  Matches   79;  Conservative    0;  Mismatches   72;  Indels    0;  Gaps    0;

Qy        489 TGGACAACAGCAACACGTGGTGTCGGGGTCTTCTGGACAACAACCGCAACAGGGAGCACA 548
              | | |  || |  |||||||   |  |  |  |   ||  |||     || ||||  ||
Db    1971184 TCGGCGTCACCTTCACGTGGCAGCTTGAACCACCACACGTCAAAGTTCACCGGGAATTCA 1971125

Qy        549 GTCAAGCACTGTCCAGCCAACAACCGGATCACCGCCCGCGGCCCAAGGCGTGCCACAGTC 608
                |     | ||||||||     ||||  | ||| ||  | ||  ||  ||||| |||
Db    1971124 CGCGCCTTCAGTCCAGCCTCGTGCCGGGCGATCGACCATCGGCCGTCGCACGCCACGGTC 1971065

Qy        609 TACCCCGCCCCCAACCCAAAATACCCCCCAG 639
              |  |||||| |||  ||      ||| | |
Db    1971064 AATTCCGCCCGCAATTCACCCGGGCCCTCGG 1971034


RESULT 8
US-09-103-840A-1/c
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
;  APPLICANT: FLEISCHMAN, Robert D.
;  APPLICANT:  WHITE, Owen R.
;  APPLICANT:  FRASER, Claire M.
;  APPLICANT:  VENTER, John C.
;  TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
;  TITLE OF INVENTION:  TUBERCULOSIS
;  FILE REFERENCE: 24366-20007.00
;  CURRENT APPLICATION NUMBER: US/09/103,840A
;  CURRENT FILING DATE:  1998-06-24
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 4411529
;   TYPE: DNA
;   ORGANISM: Mycobacterium tuberculosis
;   OTHER INFORMATION: H37Rv
US-09-103-840A-1

  Query Match             5.1%;  Score 35.8;  DB 3;  Length 4411529;
  Best Local Similarity   52.3%;  Pred. No. 15;
  Matches   79;  Conservative    0;  Mismatches   72;  Indels    0;  Gaps    0;
```

```
Qy         489 TGGACAACAGCAACACGTGGTGTCGGGGTCTTCTGGACAACAACCGCAACAGGGAGCACA 548
               | | |  || |  |||||||   |  |  |  |   ||  |||     || ||||   ||
Db     1980355 TCGGCGTCACCTTCACGTGGCAGCTTGAACCACCACACGTCAAAGTTCACCGGGAATTCA 1980296

Qy         549 GTCAAGCACTGTCCAGCCAACAACCGGATCACCGCCCGCGGCCCAAGGCGTGCCACAGTC 608
                 |     | |||||||||     ||||   | || ||   | ||   ||  ||||| |||
Db     1980295 CGCGCCTTCAGTCCAGCCTCGTGCCGGGCGATCGACCATCGGCCGTCGCACGCCACGGTC 1980236

Qy         609 TACCCCGCCCCCAACCCAAAATACCCCCCAG 639
               | |||||| |||  ||       ||| | |
Db     1980235 AATTCCGCCCGCAATTCACCCGGGCCCTCGG 1980205
```

RESULT 9
US-09-949-016-160829/c
; Sequence 160829, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 160829
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-160829

Query Match 5.0%; Score 35; DB 4; Length 601;
Best Local Similarity 50.9%; Pred. No. 0.25;
Matches 83; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

```
Qy         516 GTCTTCTGGACAACAACCGCAACAGGGAGCACAGTCAAGCACTGTCCAGCCAACAACCGG 575
               | ||||||||  || | |||| |  |||| |||||  |   |  || ||    |||  |
Db         503 GGCTTCTGGAGGACTAACGCACCCTGGAGTACAGTGACTTAAAGTGCAAGGCACAGGCCT 444

Qy         576 ATCACCGCCCGCGGCCCAAGGCGTGCCACAGTCTACCCCGCCCCCAACCCAAAATACCCC 635
               | ||| |   | | || |   |  | | | ||| |  | | | |   ||     | | |
Db         443 AGCACAGAGGGGGCCCTACCCCCAGTCTCTGTCAAGACGGACACAGCCCTTGCCTTCTCT 384

Qy         636 CCAGGGGGGTAAGGGACAGACCTTGTCACACACGGGACAATCT 678
               ||||  |   | |      ||| || |   ||   |||   ||
Db         383 CCAGCCGCCCATGATCTCCACCCTGGCCTCCAACAGACCTGCT 341
```

RESULT 10
US-09-949-016-16243
; Sequence 16243, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16243
; LENGTH: 45427
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-16243

Query Match 5.0%; Score 35; DB 4; Length 45427;
Best Local Similarity 50.9%; Pred. No. 2.5;
Matches 83; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

```
Qy         516 GTCTTCTGGACAACAACCGCAACAGGGAGCACAGTCAAGCACTGTCCAGCCAACAACCGG 575
               | ||||||||  || | |||| |  |||| |||||  |   |  || ||    |||  |
Db       34018 GGCTTCTGGAGGACTAACGCACCCTGGAGTACAGTGACTTAAAGTGCAAGGCACAGGCCT 34077

Qy         576 ATCACCGCCCGCGGCCCAAGGCGTGCCACAGTCTACCCCGCCCCCAACCCAAAATACCCC 635
               | ||| |   | | || |   |  | | | ||| |  | | | |   ||     | | |
Db       34078 AGCACAGAGGGGGCCCTACCCCCAGTCTCTGTCAAGACGGACACAGCCCTTGCCTTCTCT 34137

Qy         636 CCAGGGGGGTAAGGGACAGACCTTGTCACACACGGGACAATCT 678
               ||||  |   | |      ||| || |   ||   |||   ||
Db       34138 CCAGCCGCCCATGATCTCCACCCTGGCCTCCAACAGACCTGCT 34180
```

RESULT 11
US-08-911-853-12
; Sequence 12, Application US/08911853
; Patent No. 6048710
; GENERAL INFORMATION:
; APPLICANT: Gerritse, Gijsbert
; APPLICANT: Quax, Wilhelmus J.
; TITLE OF INVENTION: EXPRESSION SYSTEM FOR ALTERED
; TITLE OF INVENTION: EXPRESSION LEVELS
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genencor International
; STREET: 925 Page Mill Road
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1013
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/911,853
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/699,092
; FILING DATE: 16-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Glaister, Debra J
; REGISTRATION NUMBER: 33,888
; REFERENCE/DOCKET NUMBER: GC361-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-846-7620
; TELEFAX: 650-845-6504
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 642 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-911-853-12

Query Match 4.8%; Score 34; DB 3; Length 642;
Best Local Similarity 48.0%; Pred. No. 0.56;
Matches 97; Conservative 0; Mismatches 105; Indels 0; Gaps 0;

```
Qy         452 CGCAACACCCCCCACCTCAAACGACGTCGGCATCCGTTGGACAACAGCAACACGTGGTGT 511
               |||  |  | |    | || | || |  | |||||  ||| | ||    |    |  ||
```

```
Db        175 CGCCAGCCGCTGTCCATCCAGCGCCTGCAGCATCTGTTCGGCACGCCCGCGGCCAGGCCG 234

Qy        512 CGGGGTCTTCTGGACAACAACCGCAACAGGGAGCACAGTCAAGCACTGTCCAGCCAACAA 571
              || ||    | || |  |  |  || | || |||| |        ||| |||||
Db        235 CGCGGCGACCAGGCCGCCCCCGCCACCCGGCAGCAGATGACCCTGCTGGCCAGCTTCGTC 294

Qy        572 CCGGATCACCGCCCGCGGCCCAAGGCGTGCCACAGTCTACCCCGCCCCCAACCCAAAATA 631
                  ||  |  |||  ||| ||||  |  |     |  || ||  | |||||||
Db        295 AACCCGGACGCCAAGCGCTCCACGGCGATCATCCAGGTCGCCGGCGACAAACCCAAGCGC 354

Qy        632 CCCCCCAGGGGGGTAAGGGACA 653
                | ||  ||| | |  || ||
Db        355 ATCGCCGTGGGCGAATCGGTCA 376


RESULT 12
US-09-479-409-12
; Sequence 12, Application US/09479409
; Patent No. 6225106
;  GENERAL INFORMATION:
;    APPLICANT:  Gerritse, Gijsbert
;    APPLICANT:  Quax, Wilhelmus J.
;    TITLE OF INVENTION:  EXPRESSION SYSTEM FOR ALTERED
;    TITLE OF INVENTION:  EXPRESSION LEVELS
;    NUMBER OF SEQUENCES:  37
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genencor International
;      STREET:  925 Page Mill Road
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304-1013
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/479,409
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/911,853
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Glaister, Debra J
;      REGISTRATION NUMBER:  33,888
;      REFERENCE/DOCKET NUMBER:  GC361-2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-846-7620
;      TELEFAX:  650-845-6504
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  642 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-479-409-12

  Query Match             4.8%;  Score 34;  DB 3;  Length 642;
  Best Local Similarity   48.0%;  Pred. No. 0.56;
  Matches   97;  Conservative    0;  Mismatches  105;  Indels    0;  Gaps    0;

Qy        452 CGCAACACCCCCCACCTCAAACGACGTCGGCATCCGTTGGACAACAGCAACACGTGGTGT 511
              ||| |  | |    | || | || |  | ||||| ||| | ||    |    |  ||
Db        175 CGCCAGCCGCTGTCCATCCAGCGCCTGCAGCATCTGTTCGGCACGCCCGCGGCCAGGCCG 234

Qy        512 CGGGGTCTTCTGGACAACAACCGCAACAGGGAGCACAGTCAAGCACTGTCCAGCCAACAA 571
              || ||    | || |  |  |  || | || |||| |        ||| |||||
Db        235 CGCGGCGACCAGGCCGCCCCCGCCACCCGGCAGCAGATGACCCTGCTGGCCAGCTTCGTC 294

Qy        572 CCGGATCACCGCCCGCGGCCCAAGGCGTGCCACAGTCTACCCCGCCCCCAACCCAAAATA 631
                  ||  |  |||  ||| ||||  |  |     |  || ||  | |||||||
Db        295 AACCCGGACGCCAAGCGCTCCACGGCGATCATCCAGGTCGCCGGCGACAAACCCAAGCGC 354

Qy        632 CCCCCCAGGGGGGTAAGGGACA 653
                | ||  ||| | |  || ||
Db        355 ATCGCCGTGGGCGAATCGGTCA 376


RESULT 13
US-09-479-453-12
; Sequence 12, Application US/09479453
; Patent No. 6313283
;  GENERAL INFORMATION:
;    APPLICANT:  Gerritse, Gijsbert
;    APPLICANT:  Quax, Wilhelmus J.
;    TITLE OF INVENTION:  EXPRESSION SYSTEM FOR ALTERED
;    TITLE OF INVENTION:  EXPRESSION LEVELS
;    NUMBER OF SEQUENCES:  37
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genencor International
;      STREET:  925 Page Mill Road
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304-1013
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/479,453
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/911,853
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Glaister, Debra J
;      REGISTRATION NUMBER:  33,888
;      REFERENCE/DOCKET NUMBER:  GC361-2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-846-7620
;      TELEFAX:  650-845-6504
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  642 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-479-453-12

  Query Match             4.8%;  Score 34;  DB 3;  Length 642;
  Best Local Similarity   48.0%;  Pred. No. 0.56;
  Matches   97;  Conservative    0;  Mismatches  105;  Indels    0;  Gaps    0;

Qy        452 CGCAACACCCCCCACCTCAAACGACGTCGGCATCCGTTGGACAACAGCAACACGTGGTGT 511
              ||| |  | |    | || | || |  | ||||| ||| | ||    |    |  ||
Db        175 CGCCAGCCGCTGTCCATCCAGCGCCTGCAGCATCTGTTCGGCACGCCCGCGGCCAGGCCG 234

Qy        512 CGGGGTCTTCTGGACAACAACCGCAACAGGGAGCACAGTCAAGCACTGTCCAGCCAACAA 571
              || ||    | || |  |  |  || | || |||| |        ||| |||||
Db        235 CGCGGCGACCAGGCCGCCCCCGCCACCCGGCAGCAGATGACCCTGCTGGCCAGCTTCGTC 294

Qy        572 CCGGATCACCGCCCGCGGCCCAAGGCGTGCCACAGTCTACCCCGCCCCCAACCCAAAATA 631
                  ||  |  |||  ||| ||||  |  |     |  || ||  | |||||||
Db        295 AACCCGGACGCCAAGCGCTCCACGGCGATCATCCAGGTCGCCGGCGACAAACCCAAGCGC 354

Qy        632 CCCCCCAGGGGGGTAAGGGACA 653
                | ||  ||| | |  || ||
Db        355 ATCGCCGTGGGCGAATCGGTCA 376


RESULT 14
US-08-911-853-29/c
; Sequence 29, Application US/08911853
```

```
; Patent No. 6048710
;  GENERAL INFORMATION:
;    APPLICANT:  Gerritse, Gijsbert
;    APPLICANT:  Quax, Wilhelmus J.
;    TITLE OF INVENTION:  EXPRESSION SYSTEM FOR ALTERED
;    TITLE OF INVENTION:  EXPRESSION LEVELS
;    NUMBER OF SEQUENCES:  37
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genencor International
;      STREET:  925 Page Mill Road
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304-1013
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/911,853
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/699,092
;      FILING DATE:  16-AUG-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Glaister, Debra J
;      REGISTRATION NUMBER:  33,888
;      REFERENCE/DOCKET NUMBER:  GC361-2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-846-7620
;      TELEFAX:  650-845-6504
;  INFORMATION FOR SEQ ID NO:  29:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  17612 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-911-853-29

  Query Match             4.8%;  Score 34;  DB 3;  Length 17612;
  Best Local Similarity   48.0%;  Pred. No. 3.3;
  Matches   97;  Conservative    0;  Mismatches  105;  Indels    0;  Gaps    0;

Qy        452 CGCAACACCCCCCACCTCAAACGACGTCGGCATCCGTTGGACAACAGCAACACGTGGTGT 511
              ||| |  | |     | || | || |  | ||||| ||| | ||     |     |  ||
Db       3977 CGCCAGCCGCTGTCCATCCAGCGCCTGCAGCATCTGTTCGGCACGCCCGCGGCCAGGCCG 3918

Qy        512 CGGGGTCTTCTGGACAACAACCGCAACAGGGAGCACAGTCAAGCACTGTCCAGCCAACAA 571
              || ||    | || |  |  |  || | || |||| |        ||| |||||
Db       3917 CGCGGCGACCAGGCCGCCCCCGCCACCCGGCAGCAGATGACCCTGCTGGCCAGCTTCGTC 3858

Qy        572 CCGGATCACCGCCCGCGGCCCAAGGCGTGCCACAGTCTACCCCGCCCCCAACCCAAAATA 631
                 ||  |  |||  ||| ||||  |  |     |  || ||  | |||||||
Db       3857 AACCCGGACGCCAAGCGCTCCACGGCGATCATCCAGGTCGCCGGCGACAAACCCAAGCGC 3798

Qy        632 CCCCCCAGGGGGGTAAGGGACA 653
                | ||  ||| | |  || ||
Db       3797 ATCGCCGTGGGCGAATCGGTCA 3776


RESULT 15
US-09-479-409-29/c
; Sequence 29, Application US/09479409
; Patent No. 6225106
;  GENERAL INFORMATION:
;    APPLICANT:  Gerritse, Gijsbert
;    APPLICANT:  Quax, Wilhelmus J.
;    TITLE OF INVENTION:  EXPRESSION SYSTEM FOR ALTERED
;    TITLE OF INVENTION:  EXPRESSION LEVELS
;    NUMBER OF SEQUENCES:  37
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genencor International
;      STREET:  925 Page Mill Road
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304-1013
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/479,409
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/911,853
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Glaister, Debra J
;      REGISTRATION NUMBER:  33,888
;      REFERENCE/DOCKET NUMBER:  GC361-2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-846-7620
;      TELEFAX:  650-845-6504
;  INFORMATION FOR SEQ ID NO:  29:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  17612 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-479-409-29

  Query Match             4.8%;  Score 34;  DB 3;  Length 17612;
  Best Local Similarity   48.0%;  Pred. No. 3.3;
  Matches   97;  Conservative    0;  Mismatches  105;  Indels    0;  Gaps    0;

Qy        452 CGCAACACCCCCCACCTCAAACGACGTCGGCATCCGTTGGACAACAGCAACACGTGGTGT 511
              ||| |  | |     | || | || |  | ||||| ||| | ||     |     |  ||
Db       3977 CGCCAGCCGCTGTCCATCCAGCGCCTGCAGCATCTGTTCGGCACGCCCGCGGCCAGGCCG 3918

Qy        512 CGGGGTCTTCTGGACAACAACCGCAACAGGGAGCACAGTCAAGCACTGTCCAGCCAACAA 571
              || ||    | || |  |  |  || | || |||| |        ||| |||||
Db       3917 CGCGGCGACCAGGCCGCCCCCGCCACCCGGCAGCAGATGACCCTGCTGGCCAGCTTCGTC 3858

Qy        572 CCGGATCACCGCCCGCGGCCCAAGGCGTGCCACAGTCTACCCCGCCCCCAACCCAAAATA 631
                 ||  |  |||  ||| ||||  |  |     |  || ||  | |||||||
Db       3857 AACCCGGACGCCAAGCGCTCCACGGCGATCATCCAGGTCGCCGGCGACAAACCCAAGCGC 3798

Qy        632 CCCCCCAGGGGGGTAAGGGACA 653
                | ||  ||| | |  || ||
Db       3797 ATCGCCGTGGGCGAATCGGTCA 3776


Search completed: September 25, 2005, 23:49:06
Job time : 207 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: September 25, 2005, 22:52:03 ; Search time 676 Seconds
(without alignments)
6972.991 Million cell updates/sec

Title: US-09-874-140-1
Perfect score: 705
Sequence: 1 atgacacaacccgcatcgtc.....cttcaagaagtcgtagggtg 705

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 7400704 seqs, 3343079526 residues

Total number of hits satisfying chosen parameters: 14801408

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq:*
20: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq:*
21: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq:*
22: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
23: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq:*
24: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
25: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
26: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	705	100.0	705	12	US-09-874-140-1	Sequence 1, Appli
c 2	705	100.0	124884	10	US-09-913-514-1	Sequence 1, Appli
c 3	705	100.0	124884	16	US-10-288-823-76	Sequence 76, Appl
c 4	703.4	99.8	125157	10	US-09-913-514-2	Sequence 2, Appli
c 5	38	5.4	594	14	US-10-123-155-10	Sequence 10, Appl
c 6	38	5.4	594	15	US-10-146-731-10	Sequence 10, Appl
c 7	38	5.4	594	15	US-10-140-472-10	Sequence 10, Appl
c 8	38	5.4	594	15	US-10-141-761-10	Sequence 10, Appl
c 9	38	5.4	594	16	US-10-142-885-10	Sequence 10, Appl
c 10	38	5.4	594	16	US-10-158-790-10	Sequence 10, Appl
c 11	38	5.4	594	17	US-10-137-871-10	Sequence 10, Appl
c 12	38	5.4	594	17	US-10-140-923-10	Sequence 10, Appl
c 13	38	5.4	594	17	US-10-141-756-10	Sequence 10, Appl
c 14	38	5.4	594	17	US-10-141-759-10	Sequence 10, Appl
c 15	38	5.4	594	17	US-10-140-805-10	Sequence 10, Appl
c 16	38	5.4	594	17	US-10-140-864-10	Sequence 10, Appl
c 17	38	5.4	594	18	US-10-142-426-10	Sequence 10, Appl
c 18	37.8	5.4	145597	19	US-10-624-149A-2	Sequence 2, Appli
c 19	37	5.2	711	17	US-10-282-122A-31860	Sequence 31860, A
c 20	36.6	5.2	573	19	US-10-437-963-9159	Sequence 9159, Ap
c 21	36.4	5.2	1049	14	US-10-123-155-358	Sequence 358, App
c 22	36.4	5.2	1049	15	US-10-146-731-358	Sequence 358, App
c 23	36.4	5.2	1049	15	US-10-140-472-358	Sequence 358, App
c 24	36.4	5.2	1049	15	US-10-141-761-358	Sequence 358, App
c 25	36.4	5.2	1049	16	US-10-142-885-358	Sequence 358, App
c 26	36.4	5.2	1049	16	US-10-158-790-358	Sequence 358, App
c 27	36.4	5.2	1049	17	US-10-137-871-358	Sequence 358, App
c 28	36.4	5.2	1049	17	US-10-140-923-358	Sequence 358, App
c 29	36.4	5.2	1049	17	US-10-141-756-358	Sequence 358, App
c 30	36.4	5.2	1049	17	US-10-141-759-358	Sequence 358, App
c 31	36.4	5.2	1049	17	US-10-140-805-358	Sequence 358, App
c 32	36.4	5.2	1049	17	US-10-140-864-358	Sequence 358, App
c 33	36.4	5.2	1049	18	US-10-142-426-358	Sequence 358, App
34	36.2	5.1	843	19	US-10-437-963-98473	Sequence 98473, A
c 35	35.8	5.1	1236	17	US-10-282-122A-26326	Sequence 26326, A
c 36	35.8	5.1	1383	17	US-10-282-122A-28420	Sequence 28420, A
37	35.6	5.0	1222	20	US-10-335-053-193	Sequence 193, App
c 38	35.6	5.0	13416	15	US-10-205-032-11	Sequence 11, Appl
39	35.6	5.0	28874	22	US-10-915-740A-52	Sequence 52, Appl
c 40	35.6	5.0	60196	15	US-10-205-032-1	Sequence 1, Appli
41	35.6	5.0	2242716	22	US-10-915-740A-1068	Sequence 1068, Ap
c 42	35.4	5.0	646	20	US-10-363-345A-33541	Sequence 33541, A
43	35.4	5.0	646	20	US-10-363-345A-33542	Sequence 33542, A
c 44	35.4	5.0	646	21	US-10-363-483A-33541	Sequence 33541, A
45	35.4	5.0	646	21	US-10-363-483A-33542	Sequence 33542, A

ALIGNMENTS

```
RESULT 1
US-09-874-140-1
; Sequence 1, Application US/09874140
; Publication No. US20050202413A1
; GENERAL INFORMATION:
;  APPLICANT: EICKMANN, Markus
;  APPLICANT:  GICKLHORN, Dorothee
;  APPLICANT:  RADSAK, Klaus
;  APPLICANT:  HAUSER, Hans-Peter
;  APPLICANT:  GIESENDORF, Bernhard
;  TITLE OF INVENTION: VARICELLA ZOSTER VIRUS (VZV) IMMUNOREACTIVE PROTEIN
;  TITLE OF INVENTION:  VP26 AND ITS DIAGNOSTIC USE
;  FILE REFERENCE: 058315/0127
;  CURRENT APPLICATION NUMBER: US/09/874,140
;  CURRENT FILING DATE:  2001-06-06
;  PRIOR APPLICATION NUMBER: US/09/219,337
;  PRIOR FILING DATE: 1998-12-23
;  PRIOR APPLICATION NUMBER: DE 197 57 765.2
;  PRIOR FILING DATE: 1997-12-23
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 705
;   TYPE: DNA
;   ORGANISM: Varicella Zoster Virus
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(705)
US-09-874-140-1
```

```
  Query Match           100.0%;  Score 705;  DB 12;  Length 705;
  Best Local Similarity  100.0%;  Pred. No. 1.3e-229;
  Matches  705;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGACACAACCCGCATCGTCTCGTGTAGTCTTTGATCCCAGCAACCCCACCACATTTTCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGACACAACCCGCATCGTCTCGTGTAGTCTTTGATCCCAGCAACCCCACCACATTTTCG 60

Qy         61 GTGGAAGCAATTGCGGCTTACACCCCCGTTGCTTTAATACGACTTTTAAACGCCAGTGGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTGGAAGCAATTGCGGCTTACACCCCCGTTGCTTTAATACGACTTTTAAACGCCAGTGGA 120

Qy        121 CCTTTGCAACCTGGTCACCGTGTGGACATCGCTGATGCCAGAAGCATTTACACCGTGGGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCTTTGCAACCTGGTCACCGTGTGGACATCGCTGATGCCAGAAGCATTTACACCGTGGGA 180

Qy        181 GCCGCGGCCAGTGCCGCGCGTGCACGCGCTAACCATAATGCAAATACGATACGCCGAACG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCCGCGGCCAGTGCCGCGCGTGCACGCGCTAACCATAATGCAAATACGATACGCCGAACG 240

Qy        241 GCCATGTTTGCCGAGACTGACCCTATGACATGGTTAAGACCAACGGTTGGCTTAAAACGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCCATGTTTGCCGAGACTGACCCTATGACATGGTTAAGACCAACGGTTGGCTTAAAACGT 300

Qy        301 ACGTTTAACCCGCGTATTATACGACCACAACCCCCAAATCCATCCATGAGTTTGGGAATC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ACGTTTAACCCGCGTATTATACGACCACAACCCCCAAATCCATCCATGAGTTTGGGAATC 360

Qy        361 TCGGGGCCTACTATATTGCCGCAAAAAACACAGAGCGCCGATCAGTCTGCTTTACAACAG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TCGGGGCCTACTATATTGCCGCAAAAAACACAGAGCGCCGATCAGTCTGCTTTACAACAG 420

Qy        421 CCCGCCGCGTTGGCGTTTTCGGGATCATCCCCGCAACACCCCCCACCTCAAACGACGTCG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CCCGCCGCGTTGGCGTTTTCGGGATCATCCCCGCAACACCCCCCACCTCAAACGACGTCG 480

Qy        481 GCATCCGTTGGACAACAGCAACACGTGGTGTCGGGGTCTTCTGGACAACAACCGCAACAG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCATCCGTTGGACAACAGCAACACGTGGTGTCGGGGTCTTCTGGACAACAACCGCAACAG 540

Qy        541 GGAGCACAGTCAAGCACTGTCCAGCCAACAACCGGATCACCGCCCGCGGCCCAAGGCGTG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GGAGCACAGTCAAGCACTGTCCAGCCAACAACCGGATCACCGCCCGCGGCCCAAGGCGTG 600

Qy        601 CCACAGTCTACCCCGCCCCCAACCCAAAATACCCCCCAGGGGGGTAAGGGACAGACCTTG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CCACAGTCTACCCCGCCCCCAACCCAAAATACCCCCCAGGGGGGTAAGGGACAGACCTTG 660

Qy        661 TCACACACGGGACAATCTGGAAACGCTTCAAGAAGTCGTAGGGTG 705
              |||||||||||||||||||||||||||||||||||||||||||||
Db        661 TCACACACGGGACAATCTGGAAACGCTTCAAGAAGTCGTAGGGTG 705


RESULT 2
US-09-913-514-1/c
; Sequence 1, Application US/09913514
; Publication No. US20030082210A1
; GENERAL INFORMATION:
;  APPLICANT: GOMI, Yasuyuki
;  APPLICANT:  SUNAMACHI, Hiroki
;  APPLICANT:  TAKAHASHI, Michiaki
;  APPLICANT:  YAMANISHI, Koichi
;  TITLE OF INVENTION: Method for Quality Control of an Attenuated Varicella Live Vaccin
;  FILE REFERENCE: 0216-0454P
;  CURRENT APPLICATION NUMBER: US/09/913,514
;  CURRENT FILING DATE:  2001-12-07
;  PRIOR APPLICATION NUMBER: PCT/JP01/00678
;  PRIOR FILING DATE: 2001-01-31
;  PRIOR APPLICATION NUMBER: JP 2000-62734
;  PRIOR FILING DATE: 2000-01-31
;  NUMBER OF SEQ ID NOS: 42
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;  LENGTH: 124884
;  TYPE: DNA
;  ORGANISM: Varicella virus
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: (1)..(124884)
;  OTHER INFORMATION: Dumas Strain
US-09-913-514-1

  Query Match           100.0%;  Score 705;  DB 10;  Length 124884;
  Best Local Similarity  100.0%;  Pred. No. 1e-228;
  Matches  705;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGACACAACCCGCATCGTCTCGTGTAGTCTTTGATCCCAGCAACCCCACCACATTTTCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      43138 ATGACACAACCCGCATCGTCTCGTGTAGTCTTTGATCCCAGCAACCCCACCACATTTTCG 43079

Qy         61 GTGGAAGCAATTGCGGCTTACACCCCCGTTGCTTTAATACGACTTTTAAACGCCAGTGGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      43078 GTGGAAGCAATTGCGGCTTACACCCCCGTTGCTTTAATACGACTTTTAAACGCCAGTGGA 43019

Qy        121 CCTTTGCAACCTGGTCACCGTGTGGACATCGCTGATGCCAGAAGCATTTACACCGTGGGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      43018 CCTTTGCAACCTGGTCACCGTGTGGACATCGCTGATGCCAGAAGCATTTACACCGTGGGA 42959

Qy        181 GCCGCGGCCAGTGCCGCGCGTGCACGCGCTAACCATAATGCAAATACGATACGCCGAACG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42958 GCCGCGGCCAGTGCCGCGCGTGCACGCGCTAACCATAATGCAAATACGATACGCCGAACG 42899

Qy        241 GCCATGTTTGCCGAGACTGACCCTATGACATGGTTAAGACCAACGGTTGGCTTAAAACGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42898 GCCATGTTTGCCGAGACTGACCCTATGACATGGTTAAGACCAACGGTTGGCTTAAAACGT 42839

Qy        301 ACGTTTAACCCGCGTATTATACGACCACAACCCCCAAATCCATCCATGAGTTTGGGAATC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42838 ACGTTTAACCCGCGTATTATACGACCACAACCCCCAAATCCATCCATGAGTTTGGGAATC 42779

Qy        361 TCGGGGCCTACTATATTGCCGCAAAAAACACAGAGCGCCGATCAGTCTGCTTTACAACAG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42778 TCGGGGCCTACTATATTGCCGCAAAAAACACAGAGCGCCGATCAGTCTGCTTTACAACAG 42719

Qy        421 CCCGCCGCGTTGGCGTTTTCGGGATCATCCCCGCAACACCCCCCACCTCAAACGACGTCG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42718 CCCGCCGCGTTGGCGTTTTCGGGATCATCCCCGCAACACCCCCCACCTCAAACGACGTCG 42659

Qy        481 GCATCCGTTGGACAACAGCAACACGTGGTGTCGGGGTCTTCTGGACAACAACCGCAACAG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42658 GCATCCGTTGGACAACAGCAACACGTGGTGTCGGGGTCTTCTGGACAACAACCGCAACAG 42599

Qy        541 GGAGCACAGTCAAGCACTGTCCAGCCAACAACCGGATCACCGCCCGCGGCCCAAGGCGTG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42598 GGAGCACAGTCAAGCACTGTCCAGCCAACAACCGGATCACCGCCCGCGGCCCAAGGCGTG 42539

Qy        601 CCACAGTCTACCCCGCCCCCAACCCAAAATACCCCCCAGGGGGGTAAGGGACAGACCTTG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42538 CCACAGTCTACCCCGCCCCCAACCCAAAATACCCCCCAGGGGGGTAAGGGACAGACCTTG 42479

Qy        661 TCACACACGGGACAATCTGGAAACGCTTCAAGAAGTCGTAGGGTG 705
              |||||||||||||||||||||||||||||||||||||||||||||
Db      42478 TCACACACGGGACAATCTGGAAACGCTTCAAGAAGTCGTAGGGTG 42434


RESULT 3
US-10-288-823-76/c
; Sequence 76, Application US/10288823
; Publication No. US20030166168A1
; GENERAL INFORMATION:
;  APPLICANT: Grose, Charles
;  APPLICANT:  Santos, Richard
;  TITLE OF INVENTION: VARIANT VARICELLA-ZOSTER VIRUSES AND METHODS OF USE
;  FILE REFERENCE: 140.00110102
;  CURRENT APPLICATION NUMBER: US/10/288,823
;  CURRENT FILING DATE:  2002-11-06
```

```
;  PRIOR APPLICATION NUMBER: US 09/661,596
;  PRIOR FILING DATE: 2000-09-14
;  PRIOR APPLICATION NUMBER: US 60/153,779
;  PRIOR FILING DATE: 1999-09-14
;  NUMBER OF SEQ ID NOS: 80
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 76
;   LENGTH: 124884
;   TYPE: DNA
;   ORGANISM: Varicella zoster
US-10-288-823-76

  Query Match             100.0%;  Score 705;  DB 16;  Length 124884;
  Best Local Similarity   100.0%;  Pred. No. 1e-228;
  Matches  705;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGACACAACCCGCATCGTCTCGTGTAGTCTTTGATCCCAGCAACCCCACCACATTTTCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      43138 ATGACACAACCCGCATCGTCTCGTGTAGTCTTTGATCCCAGCAACCCCACCACATTTTCG 43079

Qy         61 GTGGAAGCAATTGCGGCTTACACCCCCGTTGCTTTAATACGACTTTTAAACGCCAGTGGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      43078 GTGGAAGCAATTGCGGCTTACACCCCCGTTGCTTTAATACGACTTTTAAACGCCAGTGGA 43019

Qy        121 CCTTTGCAACCTGGTCACCGTGTGGACATCGCTGATGCCAGAAGCATTTACACCGTGGGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      43018 CCTTTGCAACCTGGTCACCGTGTGGACATCGCTGATGCCAGAAGCATTTACACCGTGGGA 42959

Qy        181 GCCGCGGCCAGTGCCGCGCGTGCACGCGCTAACCATAATGCAAATACGATACGCCGAACG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42958 GCCGCGGCCAGTGCCGCGCGTGCACGCGCTAACCATAATGCAAATACGATACGCCGAACG 42899

Qy        241 GCCATGTTTGCCGAGACTGACCCTATGACATGGTTAAGACCAACGGTTGGCTTAAAACGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42898 GCCATGTTTGCCGAGACTGACCCTATGACATGGTTAAGACCAACGGTTGGCTTAAAACGT 42839

Qy        301 ACGTTTAACCCGCGTATTATACGACCACAACCCCCAAATCCATCCATGAGTTTGGGAATC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42838 ACGTTTAACCCGCGTATTATACGACCACAACCCCCAAATCCATCCATGAGTTTGGGAATC 42779

Qy        361 TCGGGGCCTACTATATTGCCGCAAAAAACACAGAGCGCCGATCAGTCTGCTTTACAACAG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42778 TCGGGGCCTACTATATTGCCGCAAAAAACACAGAGCGCCGATCAGTCTGCTTTACAACAG 42719

Qy        421 CCCGCCGCGTTGGCGTTTTCGGGATCATCCCCGCAACACCCCCCACCTCAAACGACGTCG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42718 CCCGCCGCGTTGGCGTTTTCGGGATCATCCCCGCAACACCCCCCACCTCAAACGACGTCG 42659

Qy        481 GCATCCGTTGGACAACAGCAACACGTGGTGTCGGGGTCTTCTGGACAACAACCGCAACAG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42658 GCATCCGTTGGACAACAGCAACACGTGGTGTCGGGGTCTTCTGGACAACAACCGCAACAG 42599

Qy        541 GGAGCACAGTCAAGCACTGTCCAGCCAACAACCGGATCACCGCCCGCGGCCCAAGGCGTG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42598 GGAGCACAGTCAAGCACTGTCCAGCCAACAACCGGATCACCGCCCGCGGCCCAAGGCGTG 42539

Qy        601 CCACAGTCTACCCCGCCCCCAACCCAAAATACCCCCCAGGGGGGTAAGGGACAGACCTTG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42538 CCACAGTCTACCCCGCCCCCAACCCAAAATACCCCCCAGGGGGGTAAGGGACAGACCTTG 42479

Qy        661 TCACACACGGGACAATCTGGAAACGCTTCAAGAAGTCGTAGGGTG 705
              |||||||||||||||||||||||||||||||||||||||||||||
Db      42478 TCACACACGGGACAATCTGGAAACGCTTCAAGAAGTCGTAGGGTG 42434


RESULT 4
US-09-913-514-2/c
; Sequence 2, Application US/09913514
; Publication No. US20030082210A1
; GENERAL INFORMATION:
;  APPLICANT: GOMI, Yasuyuki
;  APPLICANT:  SUNAMACHI, Hiroki
;  APPLICANT:  TAKAHASHI, Michiaki
;  APPLICANT:  YAMANISHI, Koichi
;  TITLE OF INVENTION: Method for Quality Control of an Attenuated Varicella Live Vaccine
;  FILE REFERENCE: 0216-0454P
;  CURRENT APPLICATION NUMBER: US/09/913,514
;  CURRENT FILING DATE:  2001-12-07
;  PRIOR APPLICATION NUMBER: PCT/JP01/00678
;  PRIOR FILING DATE: 2001-01-31
;  PRIOR APPLICATION NUMBER: JP 2000-62734
;  PRIOR FILING DATE: 2000-01-31
;  NUMBER OF SEQ ID NOS: 42
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
;   LENGTH: 125157
;   TYPE: DNA
;   ORGANISM: Varicella virus
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(125157)
;   OTHER INFORMATION: Attenuated Oka strain
US-09-913-514-2

  Query Match             99.8%;  Score 703.4;  DB 10;  Length 125157;
  Best Local Similarity   99.9%;  Pred. No. 3.5e-228;
  Matches  704;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 ATGACACAACCCGCATCGTCTCGTGTAGTCTTTGATCCCAGCAACCCCACCACATTTTCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      43113 ATGACACAACCCGCATCGTCTCGTGTAGTCTTTGATCCCAGCAACCCCACCACATTTTCG 43054

Qy         61 GTGGAAGCAATTGCGGCTTACACCCCCGTTGCTTTAATACGACTTTTAAACGCCAGTGGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      43053 GTGGAAGCAATTGCGGCTTACACCCCCGTTGCTTTAATACGACTTTTAAACGCCAGTGGA 42994

Qy        121 CCTTTGCAACCTGGTCACCGTGTGGACATCGCTGATGCCAGAAGCATTTACACCGTGGGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42993 CCTTTGCAACCTGGTCACCGTGTGGACATCGCTGATGCCAGAAGCATTTACACCGTGGGA 42934

Qy        181 GCCGCGGCCAGTGCCGCGCGTGCACGCGCTAACCATAATGCAAATACGATACGCCGAACG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42933 GCCGCGGCCAGTGCCGCGCGTGCACGCGCTAACCATAATGCAAATACGATACGCCGAACG 42874

Qy        241 GCCATGTTTGCCGAGACTGACCCTATGACATGGTTAAGACCAACGGTTGGCTTAAAACGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42873 GCCATGTTTGCCGAGACTGACCCTATGACATGGTTAAGACCAACGGTTGGCTTAAAACGT 42814

Qy        301 ACGTTTAACCCGCGTATTATACGACCACAACCCCCAAATCCATCCATGAGTTTGGGAATC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42813 ACGTTTAACCCGCGTATTATACGACCACAACCCCCAAATCCATCCATGAGTTTGGGAATC 42754

Qy        361 TCGGGGCCTACTATATTGCCGCAAAAAACACAGAGCGCCGATCAGTCTGCTTTACAACAG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42753 TCGGGGCCTACTATATTGCCGCAAAAAACACAGAGCGCCGATCAGTCTGCTTTACAACAG 42694

Qy        421 CCCGCCGCGTTGGCGTTTTCGGGATCATCCCCGCAACACCCCCCACCTCAAACGACGTCG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42693 CCCGCCGCGTTGGCGTTTTCGGGATCATCCCCGCAACACCCCCCACCTCAAACGACGTCG 42634

Qy        481 GCATCCGTTGGACAACAGCAACACGTGGTGTCGGGGTCTTCTGGACAACAACCGCAACAG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42633 GCATCCGTTGGACAACAGCAACACGTGGTGTCGGGGTCTTCTGGACAACAACCGCAACAG 42574

Qy        541 GGAGCACAGTCAAGCACTGTCCAGCCAACAACCGGATCACCGCCCGCGGCCCAAGGCGTG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42573 GGAGCACAGTCAAGCACTGTCCAGCCAACAACCGGATCACCGCCCGCGGCCCAAGGCGTG 42514

Qy        601 CCACAGTCTACCCCGCCCCCAACCCAAAATACCCCCCAGGGGGGTAAGGGACAGACCTTG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42513 CCACAGTCTACCCCGCCCCCAACCCAAAATACCCCCCAGGGGGGTAAGGGACAGACCTTG 42454

Qy        661 TCACACACGGGACAATCTGGAAACGCTTCAAGAAGTCGTAGGGTG 705
              || ||||||||||||||||||||||||||||||||||||||||||
Db      42453 TCCCACACGGGACAATCTGGAAACGCTTCAAGAAGTCGTAGGGTG 42409
```

```
RESULT 5
US-10-123-155-10/c
; Sequence 10, Application US/10123155
; Publication No. US20030068794A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin P.
;  APPLICANT:  Beresini,Maureen
;  APPLICANT:  DeForge,Laura
;  APPLICANT:  Desnoyers,Luc
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Gao,Wei-Qiang
;  APPLICANT:  Gerritsen,Mary E.
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Sherwood,Steven
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stewart,Timothy A.
;  APPLICANT:  Tumas,Daniel
;  APPLICANT:  Watanabe,Colin K
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang, Zemin
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3330R1C30
;  CURRENT APPLICATION NUMBER: US/10/123,155
;  CURRENT FILING DATE:  2002-04-15
; Prior Application removed -  See Palm or File Wrapper
;  NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 10
;   LENGTH: 594
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-10-123-155-10

  Query Match             5.4%;  Score 38;  DB 14;  Length 594;
  Best Local Similarity   8.0%;  Pred. No. 0.057;
  Matches   23;  Conservative  126;  Mismatches  140;  Indels    0;  Gaps    0;

Qy        389 CACAGAGCGCCGATCAGTCTGCTTTACAACAGCCCGCCGCGTTGGCGTTTTCGGGATCAT 448
              : ::: ::  :  :: : ::::: : : :   |:| ::::   ::::: ::: :  : |
Db        517 S.SDSYSYA.SYSYS.S.SWSSYSYSSSDDY.CYCCYYRYHCSDSYSYSYYY.CRCCYYT 458

Qy        449 CCCCGCAACACCCCCCACCTCAAACGACGTCGGCATCCGTTGGACAACAGCAACACGTGG 508
               :::::: : :|||:  :|::   :: :::: :: :::::   ::  :::| : :  ::
Db        457 .SYSRYDCHYSCCCSDYYCYSYSYSRYYSYSYSYSWSYSYSYTDYCSYRRCCCYYSYSSS 398

Qy        509 TGTCGGGGTCTTCTGGACAACAACCGCAACAGGGAGCACAGTCAAGCACTGTCCAGCCAA 568
              :: : ::: : |: :: :  :  : :   |           : | : :| |
Db        397 YSSSYSSSAYSTSSSSSSSSSYYYTSTNYC.T.CC....T..MCAABCSTTTTTTTTT. 338

Qy        569 CAACCGGATCACCGCCCGCGGCCCAAGGCGTGCCACAGTCTACCCCGCCCCCAACCCAAA 628
               : || :|  | :  :|   :::: : ::  :  ::: : | :::   :    ::|
Db        337 .HSCC.SA..A.M..YC.A.SYSYSYS.SSS.S.SYMR.HRA.SHYYTRS..S.MYCY.Y 278

Qy        629 ATACCCCCCAGGGGGGTAAGGGACAGACCTTGTCACACACGGGACAATC 677
              :    :: : :   ::: :     : :: |    ::::   : :|: :|
Db        277 M.Y..YY.YYSYYCSRKTM....TMTDM.T.T..MHMY.KYB.HCHTKC 229


RESULT 6
US-10-146-731-10/c
; Sequence 10, Application US/10146731
; Publication No. US20030129692A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin P.
;  APPLICANT:  Beresini,Maureen
;  APPLICANT:  DeForge,Laura
;  APPLICANT:  Desnoyers,Luc
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Gao,Wei-Qiang
;  APPLICANT:  Gerritsen,Mary E.
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Sherwood,Steven
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stewart,Timothy A.
;  APPLICANT:  Tumas,Daniel
;  APPLICANT:  Watanabe,Colin K
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang, Zemin
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3330R1C323
;  CURRENT APPLICATION NUMBER: US/10/146,731
;  CURRENT FILING DATE:  2002-05-15
;  Prior Application removed - See File Wrapper or Palm
;  NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 10
;   LENGTH: 594
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-10-146-731-10

  Query Match             5.4%;  Score 38;  DB 15;  Length 594;
  Best Local Similarity   8.0%;  Pred. No. 0.057;
  Matches   23;  Conservative  126;  Mismatches  140;  Indels    0;  Gaps    0;

Qy        389 CACAGAGCGCCGATCAGTCTGCTTTACAACAGCCCGCCGCGTTGGCGTTTTCGGGATCAT 448
              : ::: ::  :  :: : ::::: : : :   |:| ::::   ::::: ::: :  : |
Db        517 S.SDSYSYA.SYSYS.S.SWSSYSYSSSDDY.CYCCYYRYHCSDSYSYSYYY.CRCCYYT 458

Qy        449 CCCCGCAACACCCCCCACCTCAAACGACGTCGGCATCCGTTGGACAACAGCAACACGTGG 508
               :::::: : :|||:  :|::   :: :::: :: :::::   ::  :::| : :  ::
Db        457 .SYSRYDCHYSCCCSDYYCYSYSYSRYYSYSYSYSWSYSYSYTDYCSYRRCCCYYSYSSS 398

Qy        509 TGTCGGGGTCTTCTGGACAACAACCGCAACAGGGAGCACAGTCAAGCACTGTCCAGCCAA 568
              :: : ::: : |: :: :  :  : :   |           : | : :| |
Db        397 YSSSYSSSAYSTSSSSSSSSSYYYTSTNYC.T.CC....T..MCAABCSTTTTTTTTT. 338

Qy        569 CAACCGGATCACCGCCCGCGGCCCAAGGCGTGCCACAGTCTACCCCGCCCCCAACCCAAA 628
               : || :|  | :  :|   :::: : ::  :  ::: : | :::   :    ::|
Db        337 .HSCC.SA..A.M..YC.A.SYSYSYS.SSS.S.SYMR.HRA.SHYYTRS..S.MYCY.Y 278

Qy        629 ATACCCCCCAGGGGGGTAAGGGACAGACCTTGTCACACACGGGACAATC 677
              :    :: : :   ::: :     : :: |    ::::   : :|: :|
Db        277 M.Y..YY.YYSYYCSRKTM....TMTDM.T.T..MHMY.KYB.HCHTKC 229


RESULT 7
US-10-140-472-10/c
; Sequence 10, Application US/10140472
; Publication No. US20030138888A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin P.
;  APPLICANT:  Beresini,Maureen
;  APPLICANT:  DeForge,Laura
;  APPLICANT:  Desnoyers,Luc
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Gao,Wei-Qiang
;  APPLICANT:  Gerritsen,Mary E.
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Sherwood,Steven
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stewart,Timothy A.
;  APPLICANT:  Tumas,Daniel
;  APPLICANT:  Watanabe,Colin K
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang, Zemin
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3330R1C168
```

```
;  CURRENT APPLICATION NUMBER: US/10/140,472
;  CURRENT FILING DATE:  2002-05-06
;  Prior Apploication removed - See File Wrapper or Palm
;  NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 10
;   LENGTH: 594
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-10-140-472-10

  Query Match             5.4%;  Score 38;  DB 15;  Length 594;
  Best Local Similarity   8.0%;  Pred. No. 0.057;
  Matches   23;  Conservative  126;  Mismatches  140;  Indels    0;  Gaps    0;

Qy        389 CACAGAGCGCCGATCAGTCTGCTTTACAACAGCCCGCCGCGTTGGCGTTTTCGGGATCAT 448
              : ::: ::  :  :: : ::::: : : :   |:| ::::   ::::: ::: :  : |
Db        517 S.SDSYSYA.SYSYS.S.SWSSYSYSSSDDY.CYCCYYRYHCSDSYSYSYYY.CRCCYYT 458

Qy        449 CCCCGCAACACCCCCCACCTCAAACGACGTCGGCATCCGTTGGACAACAGCAACACGTGG 508
               :::::: : :|||:  :|::   :: :::: :: :::::   ::  :::|  : :  ::
Db        457 .SYSRYDCHYSCCCSDYYCYSYSYSRYYSYSYSYSWSYSYSYTDYCSYRRCCCYYSYSSS 398

Qy        509 TGTCGGGGTCTTCTGGACAACAACCGCAACAGGGAGCACAGTCAAGCACTGTCCAGCCAA 568
              :: : ::: : |: :: :  :  : :   |            : | : :| |
Db        397 YSSSYSSSAYSTSSSSSSSSSYYYTSTNYC.T.CC....T..MCAABCSTTTTTTTTT. 338

Qy        569 CAACCGGATCACCGCCCGCGGCCCAAGGCGTGCCACAGTCTACCCCGCCCCCAACCCAAA 628
              : || :|  | :  :|   :::: : ::  :  ::: : | :::   :    ::|
Db        337 .HSCC.SA..A.M..YC.A.SYSYSYS.SSS.S.SYMR.HRA.SHYYTRS..S.MYCY.Y 278

Qy        629 ATACCCCCCAGGGGGGTAAGGGACAGACCTTGTCACACACGGGACAATC 677
              :    :: : :   ::: :     : :: |    ::::   : :|: :|
Db        277 M.Y..YY.YYSYYCSRKTM....TMTDM.T.T..MHMY.KYB.HCHTKC 229


RESULT 8
US-10-141-761-10/c
; Sequence 10, Application US/10141761
; Publication No. US20030148432A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin P.
;  APPLICANT:  Beresini,Maureen
;  APPLICANT:  DeForge,Laura
;  APPLICANT:  Desnoyers,Luc
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Gao,Wei-Qiang
;  APPLICANT:  Gerritsen,Mary E.
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Sherwood,Steven
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stewart,Timothy A.
;  APPLICANT:  Tumas,Daniel
;  APPLICANT:  Watanabe,Colin K
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang, Zemin
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3330R1C198
;  CURRENT APPLICATION NUMBER: US/10/141,761
;  CURRENT FILING DATE:  2002-05-08
;  Prior Application removed - See Palm or File Wrapper
;  NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 10
;   LENGTH: 594
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-10-141-761-10

  Query Match             5.4%;  Score 38;  DB 15;  Length 594;
  Best Local Similarity   8.0%;  Pred. No. 0.057;
  Matches   23;  Conservative  126;  Mismatches  140;  Indels    0;  Gaps    0;

Qy        389 CACAGAGCGCCGATCAGTCTGCTTTACAACAGCCCGCCGCGTTGGCGTTTTCGGGATCAT 448
              : ::: ::  :  :: : ::::: : : :   |:| ::::   ::::: ::: :  : |
Db        517 S.SDSYSYA.SYSYS.S.SWSSYSYSSSDDY.CYCCYYRYHCSDSYSYSYYY.CRCCYYT 458

Qy        449 CCCCGCAACACCCCCCACCTCAAACGACGTCGGCATCCGTTGGACAACAGCAACACGTGG 508
               :::::: : :|||:  :|::   :: :::: :: :::::   ::  :::|  : :  ::
Db        457 .SYSRYDCHYSCCCSDYYCYSYSYSRYYSYSYSYSWSYSYSYTDYCSYRRCCCYYSYSSS 398

Qy        509 TGTCGGGGTCTTCTGGACAACAACCGCAACAGGGAGCACAGTCAAGCACTGTCCAGCCAA 568
              :: : ::: : |: :: :  :  : :   |            : | : :| |
Db        397 YSSSYSSSAYSTSSSSSSSSSYYYTSTNYC.T.CC....T..MCAABCSTTTTTTTTT. 338

Qy        569 CAACCGGATCACCGCCCGCGGCCCAAGGCGTGCCACAGTCTACCCCGCCCCCAACCCAAA 628
              : || :|  | :  :|   :::: : ::  :  ::: : | :::   :    ::|
Db        337 .HSCC.SA..A.M..YC.A.SYSYSYS.SSS.S.SYMR.HRA.SHYYTRS..S.MYCY.Y 278

Qy        629 ATACCCCCCAGGGGGGTAAGGGACAGACCTTGTCACACACGGGACAATC 677
              :    :: : :   ::: :     : :: |    ::::   : :|: :|
Db        277 M.Y..YY.YYSYYCSRKTM....TMTDM.T.T..MHMY.KYB.HCHTKC 229


RESULT 9
US-10-142-885-10/c
; Sequence 10, Application US/10142885
; Publication No. US20030157604A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin P.
;  APPLICANT:  Beresini,Maureen
;  APPLICANT:  DeForge,Laura
;  APPLICANT:  Desnoyers,Luc
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Gao,Wei-Qiang
;  APPLICANT:  Gerritsen,Mary E.
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Sherwood,Steven
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stewart,Timothy A.
;  APPLICANT:  Tumas,Daniel
;  APPLICANT:  Watanabe,Colin K
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang, Zemin
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3330R1C248
;  CURRENT APPLICATION NUMBER: US/10/142,885
;  CURRENT FILING DATE:  2002-05-10
;  Prior Apploication removed - See File Wrapper or Palm
;  NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 10
;   LENGTH: 594
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-10-142-885-10

  Query Match             5.4%;  Score 38;  DB 16;  Length 594;
  Best Local Similarity   8.0%;  Pred. No. 0.057;
  Matches   23;  Conservative  126;  Mismatches  140;  Indels    0;  Gaps    0;

Qy        389 CACAGAGCGCCGATCAGTCTGCTTTACAACAGCCCGCCGCGTTGGCGTTTTCGGGATCAT 448
              : ::: ::  :  :: : ::::: : : :   |:| ::::   ::::: ::: :  : |
Db        517 S.SDSYSYA.SYSYS.S.SWSSYSYSSSDDY.CYCCYYRYHCSDSYSYSYYY.CRCCYYT 458

Qy        449 CCCCGCAACACCCCCCACCTCAAACGACGTCGGCATCCGTTGGACAACAGCAACACGTGG 508
               :::::: : :|||:  :|::   :: :::: :: :::::   ::  :::|  : :  ::
Db        457 .SYSRYDCHYSCCCSDYYCYSYSYSRYYSYSYSYSWSYSYSYTDYCSYRRCCCYYSYSSS 398

Qy        509 TGTCGGGGTCTTCTGGACAACAACCGCAACAGGGAGCACAGTCAAGCACTGTCCAGCCAA 568
              :: : ::: : |: :: :  :  : :   |            : | : :| |
Db        397 YSSSYSSSAYSTSSSSSSSSSYYYTSTNYC.T.CC....T..MCAABCSTTTTTTTTT. 338
```

```
Qy        569 CAACCGGATCACCGCCCGCGGCCCAAGGCGTGCCACAGTCTACCCCGCCCCCAACCCAAA 628
              : || :|  | :  :|   :::: : ::  :  ::: : | :::   :    ::|
Db        337 .HSCC.SA..A.M..YC.A.SYSYSYS.SSS.S.SYMR.HRA.SHYYTRS..S.MYCY.Y 278

Qy        629 ATACCCCCCAGGGGGGTAAGGGACAGACCTTGTCACACACGGGACAATC 677
              :    :: : :   ::: :     : :: |    ::::   : :|: :|
Db        277 M.Y..YY.YYSYYCSRKTM....TMTDM.T.T..MHMY.KYB.HCHTKC 229


RESULT 10
US-10-158-790-10/c
; Sequence 10, Application US/10158790
; Publication No. US20030180879A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin P.
;  APPLICANT:  Beresini,Maureen
;  APPLICANT:  DeForge,Laura
;  APPLICANT:  Desnoyers,Luc
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Gao,Wei-Qiang
;  APPLICANT:  Gerritsen,Mary E.
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Sherwood,Steven
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stewart,Timothy A.
;  APPLICANT:  Tumas,Daniel
;  APPLICANT:  Watanabe,Colin K
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang, Zemin
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3330R1C448
;  CURRENT APPLICATION NUMBER: US/10/158,790
;  CURRENT FILING DATE:  2002-05-30
;  Prior Application removed - See File Wrapper or Palm
;  NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 10
;   LENGTH: 594
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-10-158-790-10

  Query Match             5.4%;  Score 38;  DB 16;  Length 594;
  Best Local Similarity   8.0%;  Pred. No. 0.057;
  Matches   23;  Conservative  126;  Mismatches  140;  Indels    0;  Gaps    0;

Qy        389 CACAGAGCGCCGATCAGTCTGCTTTACAACAGCCCGCCGCGTTGGCGTTTTCGGGATCAT 448
              : ::: ::  :  :: : ::::: : : :   |:| ::::   ::::: ::: :  : |
Db        517 S.SDSYSYA.SYSYS.S.SWSSYSYSSSDDY.CYCCYYRYHCSDSYSYSYYY.CRCCYYT 458

Qy        449 CCCCGCAACACCCCCCACCTCAAACGACGTCGGCATCCGTTGGACAACAGCAACACGTGG 508
               :::::: : :|||:  :|::   :: :::: :: ::::::   ::  :::|  : :  ::
Db        457 .SYSRYDCHYSCCCSDYYCYSYSYSRYYSYSYSYSWSYSYSYTDYCSYRRCCCYYSYSSS 398

Qy        509 TGTCGGGGTCTTCTGGACAACAACCGCAACAGGGAGCACAGTCAAGCACTGTCCAGCCAA 568
              :: : ::: : |: :: :  :  : :   |             : | : :| |
Db        397 YSSSYSSSAYSTSSSSSSSSSYYYTSTNYC.T.CC....T..MCAABCSTTTTTTTTTT. 338

Qy        569 CAACCGGATCACCGCCCGCGGCCCAAGGCGTGCCACAGTCTACCCCGCCCCCAACCCAAA 628
              : || :|  | :  :|   :::: : ::  :  ::: : | :::   :    ::|
Db        337 .HSCC.SA..A.M..YC.A.SYSYSYS.SSS.S.SYMR.HRA.SHYYTRS..S.MYCY.Y 278

Qy        629 ATACCCCCCAGGGGGGTAAGGGACAGACCTTGTCACACACGGGACAATC 677
              :    :: : :   ::: :     : :: |    ::::   : :|: :|
Db        277 M.Y..YY.YYSYYCSRKTM....TMTDM.T.T..MHMY.KYB.HCHTKC 229


RESULT 11
US-10-137-871-10/c
; Sequence 10, Application US/10137871
; Publication No. US20030207350A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin P.
;  APPLICANT:  Beresini,Maureen
;  APPLICANT:  DeForge,Laura
;  APPLICANT:  Desnoyers,Luc
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Gao,Wei-Qiang
;  APPLICANT:  Gerritsen,Mary E.
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Sherwood,Steven
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stewart,Timothy A.
;  APPLICANT:  Tumas,Daniel
;  APPLICANT:  Watanabe,Colin K
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang, Zemin
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3330R1C153
;  CURRENT APPLICATION NUMBER: US/10/137,871
;  CURRENT FILING DATE:  2002-05-03
;  Prior Application removed - See Palm or File Wrapper
;  NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 10
;   LENGTH: 594
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-10-137-871-10

  Query Match             5.4%;  Score 38;  DB 17;  Length 594;
  Best Local Similarity   8.0%;  Pred. No. 0.057;
  Matches   23;  Conservative  126;  Mismatches  140;  Indels    0;  Gaps    0;

Qy        389 CACAGAGCGCCGATCAGTCTGCTTTACAACAGCCCGCCGCGTTGGCGTTTTCGGGATCAT 448
              : ::: ::  :  :: : ::::: : : :   |:| ::::   ::::: ::: :  : |
Db        517 S.SDSYSYA.SYSYS.S.SWSSYSYSSSDDY.CYCCYYRYHCSDSYSYSYYY.CRCCYYT 458

Qy        449 CCCCGCAACACCCCCCACCTCAAACGACGTCGGCATCCGTTGGACAACAGCAACACGTGG 508
               :::::: : :|||:  :|::   :: :::: :: ::::::   ::  :::|  : :  ::
Db        457 .SYSRYDCHYSCCCSDYYCYSYSYSRYYSYSYSYSWSYSYSYTDYCSYRRCCCYYSYSSS 398

Qy        509 TGTCGGGGTCTTCTGGACAACAACCGCAACAGGGAGCACAGTCAAGCACTGTCCAGCCAA 568
              :: : ::: : |: :: :  :  : :   |             : | : :| |
Db        397 YSSSYSSSAYSTSSSSSSSSSYYYTSTNYC.T.CC....T..MCAABCSTTTTTTTTTT. 338

Qy        569 CAACCGGATCACCGCCCGCGGCCCAAGGCGTGCCACAGTCTACCCCGCCCCCAACCCAAA 628
              : || :|  | :  :|   :::: : ::  :  ::: : | :::   :    ::|
Db        337 .HSCC.SA..A.M..YC.A.SYSYSYS.SSS.S.SYMR.HRA.SHYYTRS..S.MYCY.Y 278

Qy        629 ATACCCCCCAGGGGGGTAAGGGACAGACCTTGTCACACACGGGACAATC 677
              :    :: : :   ::: :     : :: |    ::::   : :|: :|
Db        277 M.Y..YY.YYSYYCSRKTM....TMTDM.T.T..MHMY.KYB.HCHTKC 229


RESULT 12
US-10-140-923-10/c
; Sequence 10, Application US/10140923
; Publication No. US20030207355A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin P.
;  APPLICANT:  Beresini,Maureen
;  APPLICANT:  DeForge,Laura
;  APPLICANT:  Desnoyers,Luc
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Gao,Wei-Qiang
;  APPLICANT:  Gerritsen,Mary E.
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Sherwood,Steven
;  APPLICANT:  Smith,Victoria
```

```
;  APPLICANT:  Stewart,Timothy A.
;  APPLICANT:  Tumas,Daniel
;  APPLICANT:  Watanabe,Colin K
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang, Zemin
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3330R1C188
;  CURRENT APPLICATION NUMBER: US/10/140,923
;  CURRENT FILING DATE:  2002-05-07
;  Prior Application removed - See Palm or File Wrapper
;  NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 10
;   LENGTH: 594
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-10-140-923-10

  Query Match             5.4%;  Score 38;  DB 17;  Length 594;
  Best Local Similarity   8.0%;  Pred. No. 0.057;
  Matches   23;  Conservative  126;  Mismatches  140;  Indels    0;  Gaps    0;

Qy        389 CACAGAGCGCCGATCAGTCTGCTTTACAACAGCCCGCCGCGTTGGCGTTTTCGGGATCAT 448
              : ::: ::  :  :: : ::::: : : :   |:| ::::   ::::: ::: :  : |
Db        517 S.SDSYSYA.SYSYS.S.SWSSYSYSSSDDY.CYCCYYRYHCSDSYSYSYYY.CRCCYYT 458

Qy        449 CCCCGCAACACCCCCCACCTCAAACGACGTCGGCATCCGTTGGACAACAGCAACACGTGG 508
              :::::: : :|||:  :|::   :: :::: :: :::::   ::  :::|  : :  ::
Db        457 .SYSRYDCHYSCCCSDYYCYSYSYSRYYSYSYSYSWSYSYSYTDYCSYRRCCCYYSYSSS 398

Qy        509 TGTCGGGGTCTTCTGGACAACAACCGCAACAGGGAGCACAGTCAAGCACTGTCCAGCCAA 568
              :: : ::: : |: :: :  :  : :   |             : | : :| |
Db        397 YSSSYSSSAYSTSSSSSSSSSYYYTSTNYC.T.CC....T..MCAABCSTTTTTTTTTT. 338

Qy        569 CAACCGGATCACCGCCCGCGGCCCAAGGCGTGCCACAGTCTACCCCGCCCCCAACCCAAA 628
              : || :|  | :  :|   :::  : ::  :  ::: : | :::   :    ::|
Db        337 .HSCC.SA..A.M..YC.A.SYSYSYS.SSS.S.SYMR.HRA.SHYYTRS..S.MYCY.Y 278

Qy        629 ATACCCCCCAGGGGGGTAAGGGACAGACCTTGTCACACACGGGACAATC 677
              :    :: : :   ::: :     : :: |    ::::   : :|: :|
Db        277 M.Y..YY.YYSYYCSRKTM....TMTDM.T.T..MHMY.KYB.HCHTKC 229


RESULT 13
US-10-141-756-10/c
; Sequence 10, Application US/10141756
; Publication No. US20030207359A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin P.
;  APPLICANT:  Beresini,Maureen
;  APPLICANT:  DeForge,Laura
;  APPLICANT:  Desnoyers,Luc
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Gao,Wei-Qiang
;  APPLICANT:  Gerritsen,Mary E.
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Sherwood,Steven
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stewart,Timothy A.
;  APPLICANT:  Tumas,Daniel
;  APPLICANT:  Watanabe,Colin K
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang, Zemin
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3330R1C200
;  CURRENT APPLICATION NUMBER: US/10/141,756
;  CURRENT FILING DATE:  2002-05-08
;  Prior Apploication removed - See File Wrapper or Palm
;  NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 10
;   LENGTH: 594
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-10-141-756-10

  Query Match             5.4%;  Score 38;  DB 17;  Length 594;
  Best Local Similarity   8.0%;  Pred. No. 0.057;
  Matches   23;  Conservative  126;  Mismatches  140;  Indels    0;  Gaps    0;

Qy        389 CACAGAGCGCCGATCAGTCTGCTTTACAACAGCCCGCCGCGTTGGCGTTTTCGGGATCAT 448
              : ::: ::  :  :: : ::::: : : :   |:| ::::   ::::: ::: :  : |
Db        517 S.SDSYSYA.SYSYS.S.SWSSYSYSSSDDY.CYCCYYRYHCSDSYSYSYYY.CRCCYYT 458

Qy        449 CCCCGCAACACCCCCCACCTCAAACGACGTCGGCATCCGTTGGACAACAGCAACACGTGG 508
              :::::: : :|||:  :|::   :: :::: :: :::::   ::  :::|  : :  ::
Db        457 .SYSRYDCHYSCCCSDYYCYSYSYSRYYSYSYSYSWSYSYSYTDYCSYRRCCCYYSYSSS 398

Qy        509 TGTCGGGGTCTTCTGGACAACAACCGCAACAGGGAGCACAGTCAAGCACTGTCCAGCCAA 568
              :: : ::: : |: :: :  :  : :   |             : | : :| |
Db        397 YSSSYSSSAYSTSSSSSSSSSYYYTSTNYC.T.CC....T..MCAABCSTTTTTTTTTT. 338

Qy        569 CAACCGGATCACCGCCCGCGGCCCAAGGCGTGCCACAGTCTACCCCGCCCCCAACCCAAA 628
              : || :|  | :  :|   :::  : ::  :  ::: : | :::   :    ::|
Db        337 .HSCC.SA..A.M..YC.A.SYSYSYS.SSS.S.SYMR.HRA.SHYYTRS..S.MYCY.Y 278

Qy        629 ATACCCCCCAGGGGGGTAAGGGACAGACCTTGTCACACACGGGACAATC 677
              :    :: : :   ::: :     : :: |    ::::   : :|: :|
Db        277 M.Y..YY.YYSYYCSRKTM....TMTDM.T.T..MHMY.KYB.HCHTKC 229


RESULT 14
US-10-141-759-10/c
; Sequence 10, Application US/10141759
; Publication No. US20030207361A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin P.
;  APPLICANT:  Beresini,Maureen
;  APPLICANT:  DeForge,Laura
;  APPLICANT:  Desnoyers,Luc
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Gao,Wei-Qiang
;  APPLICANT:  Gerritsen,Mary E.
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Sherwood,Steven
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stewart,Timothy A.
;  APPLICANT:  Tumas,Daniel
;  APPLICANT:  Watanabe,Colin K
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang, Zemin
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3330R1C197
;  CURRENT APPLICATION NUMBER: US/10/141,759
;  CURRENT FILING DATE:  2002-05-08
;  Prior Apploication removed - See File Wrapper or Palm
;  NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 10
;   LENGTH: 594
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-10-141-759-10

  Query Match             5.4%;  Score 38;  DB 17;  Length 594;
  Best Local Similarity   8.0%;  Pred. No. 0.057;
  Matches   23;  Conservative  126;  Mismatches  140;  Indels    0;  Gaps    0;

Qy        389 CACAGAGCGCCGATCAGTCTGCTTTACAACAGCCCGCCGCGTTGGCGTTTTCGGGATCAT 448
              : ::: ::  :  :: : ::::: : : :   |:| ::::   ::::: ::: :  : |
Db        517 S.SDSYSYA.SYSYS.S.SWSSYSYSSSDDY.CYCCYYRYHCSDSYSYSYYY.CRCCYYT 458
```

```
Qy        449 CCCCGCAACACCCCCCACCTCAAACGACGTCGGCATCCGTTGGACAACAGCAACACGTGG 508
              ::::::  : :|||:  :|::   :: :::: :: :::::   ::  :::|  : :  ::
Db        457 .SYSRYDCHYSCCCSDYYCYSYSYSRYYSYSYSYSWSYSYSYTDYCSYRRCCCYYSYSSS 398

Qy        509 TGTCGGGGTCTTCTGGACAACAACCGCAACAGGGAGCACAGTCAAGCACTGTCCAGCCAA 568
              :: : ::: : |: :: :  :  : :   |           : | : :| |
Db        397 YSSSYSSSAYSTSSSSSSSSSYYYTSTNYC.T.CC....T..MCAABCSTTTTTTTTTT. 338

Qy        569 CAACCGGATCACCGCCCGCGGCCCAAGGCGTGCCACAGTCTACCCCGCCCCCAACCCAAA 628
              : || :|  | :  :|   :::: : ::  :  ::: : | :::   :    ::|
Db        337 .HSCC.SA..A.M..YC.A.SYSYSYS.SSS.S.SYMR.HRA.SHYYTRS..S.MYCY.Y 278

Qy        629 ATACCCCCCAGGGGGGTAAGGGACAGACCTTGTCACACACGGGACAATC 677
              :   :: : :   ::: :     : :: |    ::::   : :|: :|
Db        277 M.Y..YY.YYSYYCSRKTM....TMTDM.T.T..MHMY.KYB.HCHTKC 229


RESULT 15
US-10-140-805-10/c
; Sequence 10, Application US/10140805
; Publication No. US20030207417A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin P.
;  APPLICANT:  Beresini,Maureen
;  APPLICANT:  DeForge,Laura
;  APPLICANT:  Desnoyers,Luc
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Gao,Wei-Qiang
;  APPLICANT:  Gerritsen,Mary E.
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Sherwood,Steven
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stewart,Timothy A.
;  APPLICANT:  Tumas,Daniel
;  APPLICANT:  Watanabe,Colin K
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang, Zemin
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3330R1C176
;  CURRENT APPLICATION NUMBER: US/10/140,805
;  CURRENT FILING DATE:  2002-05-07
;  Prior Apploication removed - See File Wrapper or Palm
;  NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 10
;   LENGTH: 594
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-10-140-805-10

  Query Match             5.4%;  Score 38;  DB 17;  Length 594;
  Best Local Similarity   8.0%;  Pred. No. 0.057;
  Matches   23;  Conservative  126;  Mismatches  140;  Indels    0;  Gaps    0;

Qy        389 CACAGAGCGCCGATCAGTCTGCTTTACAACAGCCCGCCGCGTTGGCGTTTTCGGGATCAT 448
              : ::: ::  :  :: : ::::: : : :   |:| ::::   ::::: :::  :  : |
Db        517 S.SDSYSYA.SYSYS.S.SWSSYSYSSSDDY.CYCCYYRYHCSDSYSYSYYY.CRCCYYT 458

Qy        449 CCCCGCAACACCCCCCACCTCAAACGACGTCGGCATCCGTTGGACAACAGCAACACGTGG 508
              ::::::  : :|||:  :|::   :: :::: :: :::::   ::  :::|  : :  ::
Db        457 .SYSRYDCHYSCCCSDYYCYSYSYSRYYSYSYSYSWSYSYSYTDYCSYRRCCCYYSYSSS 398

Qy        509 TGTCGGGGTCTTCTGGACAACAACCGCAACAGGGAGCACAGTCAAGCACTGTCCAGCCAA 568
              :: : ::: : |: :: :  :  : :   |           : | : :| |
Db        397 YSSSYSSSAYSTSSSSSSSSSYYYTSTNYC.T.CC....T..MCAABCSTTTTTTTTTT. 338

Qy        569 CAACCGGATCACCGCCCGCGGCCCAAGGCGTGCCACAGTCTACCCCGCCCCCAACCCAAA 628
              : || :|  | :  :|   :::: : ::  :  ::: : | :::   :    ::|
Db        337 .HSCC.SA..A.M..YC.A.SYSYSYS.SSS.S.SYMR.HRA.SHYYTRS..S.MYCY.Y 278

Qy        629 ATACCCCCCAGGGGGGTAAGGGACAGACCTTGTCACACACGGGACAATC 677
              :   :: : :   ::: :     : :: |    ::::   : :|: :|
Db        277 M.Y..YY.YYSYYCSRKTM....TMTDM.T.T..MHMY.KYB.HCHTKC 229


Search completed: September 26, 2005, 01:30:26
Job time : 679 secs
```